

# TWIN PEAKS

KRZYSZTOF BURDZY AND SOUMIK PAL

**ABSTRACT.** We study random labelings of graphs conditioned on a small number (typically one or two) peaks, i.e., local maxima. We show that the boundaries of level sets of a random labeling of a square with a single peak have dimension 2, in a suitable asymptotic sense. The gradient line of a random labeling of a long ladder graph conditioned on a single peak consists mostly of straight line segments. We show that for some tree-graphs, if a random labeling is conditioned on exactly two peaks then the peaks can be very close to each other. We also study random labelings of regular trees conditioned on having exactly two peaks. Our results suggest that the top peak is likely to be at the root and the second peak is equally likely, more or less, to be any vertex not adjacent to the root.

## 1. INTRODUCTION

This paper is the first part of a project motivated by our desire to understand random labelings of graphs conditioned on having a small number of peaks (local maxima). The roots of this project lie in the paper [BBS13], where some new results on peaks of random permutations were proved. These were later applied in [BBPS15].

To indicate the direction of our thinking, we now state informally an open problem that was the target of our initial investigation. Consider a discrete 2-dimensional torus with  $N$  vertices and let  $L$  be a random labeling of the vertices of the torus with integers  $1, 2, \dots, N$ . If we condition  $L$  on having exactly two peaks and we call their locations  $K_1$  and  $K_2$ , is it true that  $\text{dist}(K_1, K_2)/\sqrt{N} \rightarrow 0$  in probability as  $N \rightarrow \infty$ ? We conjecture that the answer is “no.”

We are grateful to Nicolas Curien for pointing out to us the following related articles. Local maxima of labelings are analyzed in [AB13]. The labels (ages) of vertices in a random tree generated according to the Barabási and Albert “preferential attachment model” (see [VdH14, Chap. 8]), conditioned on a given tree structure, have the distribution of a random labeling conditioned on having a single peak at the root. These research directions are significantly different from ours.

The paper is organized as follows. It starts with Section 2 presenting a comparison between a 1-dimensional model and a 2-dimensional model, and another comparison of two 2-dimensional models, as a further motivation for our study. Section 3 contains some preliminaries (notation and basic formulas). In Sections 4 and 5 we will prove two results on random labelings of tori inspired by simulations. Section 6 will present results on random labelings of trees. One can prove more precise results for random labelings of *regular* trees—these will be given in Section 7.

---

Research supported in part by NSF Grants DMS-1206276 for Burdzy and DMS-1308340 for Pal.

## 2. MODELS AND SIMULATIONS

We will define a random labeling of  $N$  vertices of  $\mathbb{Z}^d$  and a random labeling of  $d$ -dimensional torus with  $N$  vertices. The two algorithms generate similarly looking random labelings in dimension  $d = 1$ , but computer simulations and some rigorous results (to be presented later in this paper) show that the two algorithms generate strikingly different random labelings in  $\mathbb{Z}^2$ .

(i) Fix an integer  $N > 0$ . We will define inductively a random function  $\tilde{L} : \mathbb{Z}^d \rightarrow \{0, 1, 2, \dots, N\}$  such that the origin is mapped onto  $N$  and  $\tilde{L}^{-1}(j)$  is a singleton for  $1 \leq j \leq N$ . Let  $\tilde{C}_k$  be the inverse image of  $\{N, N-1, \dots, N-k\}$  by  $\tilde{L}$  for  $0 \leq k \leq N-1$ . Given  $\tilde{C}_k$  for some  $0 \leq k \leq N-2$ , the vertex  $\tilde{L}^{-1}(N-k-1)$  is distributed uniformly among all vertices in  $\mathbb{Z}^d \setminus \tilde{C}_k$  which have a neighbor in  $\tilde{C}_k$ . We let  $\tilde{L}(x) = 0$  for all  $x \notin \tilde{C}_{N-1}$ . It is easy to see that the only peak (strict local maximum) of  $\tilde{L}$  is at  $\tilde{L}^{-1}(N)$ .

(ii) Suppose that  $N_1 > 0$ ,  $N = N_1^d$  and let  $\mathcal{G}$  be the  $d$ -dimensional torus with side length  $N_1$  and vertex set  $\mathcal{V}$ . Let  $L : \mathcal{V} \rightarrow \{1, 2, \dots, N\}$  be a random (uniformly chosen) bijection conditioned on  $N$  being the only peak. Let  $C_k$  be the inverse image of  $\{N, N-1, \dots, N-k\}$  by  $L$  for  $0 \leq k \leq N-1$ .

Note that both  $\tilde{C}_k$  and  $C_k$  are connected for every  $0 \leq k \leq N-1$ . We also have  $\tilde{C}_k \subset \tilde{C}_{k+1}$  and  $C_k \subset C_{k+1}$  for all  $k$ .

Methods applied in [BBS13] and other standard techniques show that in dimension  $d = 1$ , for large  $N$  and  $k$  comparable with  $N$ , both  $\tilde{C}_k$  and  $C_k$  are intervals centered, more or less, at the vertex labeled  $N$ , with high probability.

In dimension  $d = 2$ , the cluster growth process  $\{\tilde{C}_k, k \geq 1\}$  is the Eden model, [Ede61]. See Fig. 1. Convincing heuristic arguments (see Remark 4.1) show that the boundary of  $\tilde{C}_k$  has the size of the order  $k^{1/2}$ .

The contrast between Figs. 1 and 2 could not be more dramatic. Fig. 2 is a simulation of the process  $\{C_k, k \geq 1\}$  in two dimensions. We will show that the boundaries of the sets  $C_k$  have dimension 2, in an asymptotic sense.

The following elementary example shows that models (i) and (ii) are not equivalent for general graphs. In other words, the sets  $C_k$  associated with the (uniformly chosen) random labeling with a single peak do not grow by attaching vertices uniformly on the boundary.

**Example 2.1.** Consider a graph with vertices  $\{a, b, c, d\}$  and edges  $\{\overline{ab}, \overline{bc}, \overline{cd}\}$  (a linear graph, with  $a, b, c$  and  $d$  arranged along a line in this order). Let  $L$  be the uniformly random labeling conditioned on having a single peak at  $b$ . Let  $R = L^{-1}$  be the inverse function of  $L$ . The following three realizations of  $L$  are possible and equally likely.

$$\begin{aligned} L(a) = 3, L(b) = 4, L(c) = 2, L(d) = 1; \\ L(a) = 2, L(b) = 4, L(c) = 3, L(d) = 1; \\ L(a) = 1, L(b) = 4, L(c) = 3, L(d) = 2. \end{aligned}$$

We see that  $R(3)$  can be equal to  $a$  or  $c$  but

$$1/3 = \mathbb{P}(R(3) = a \mid R(4) = b) \neq \mathbb{P}(R(3) = c \mid R(4) = b) = 2/3.$$

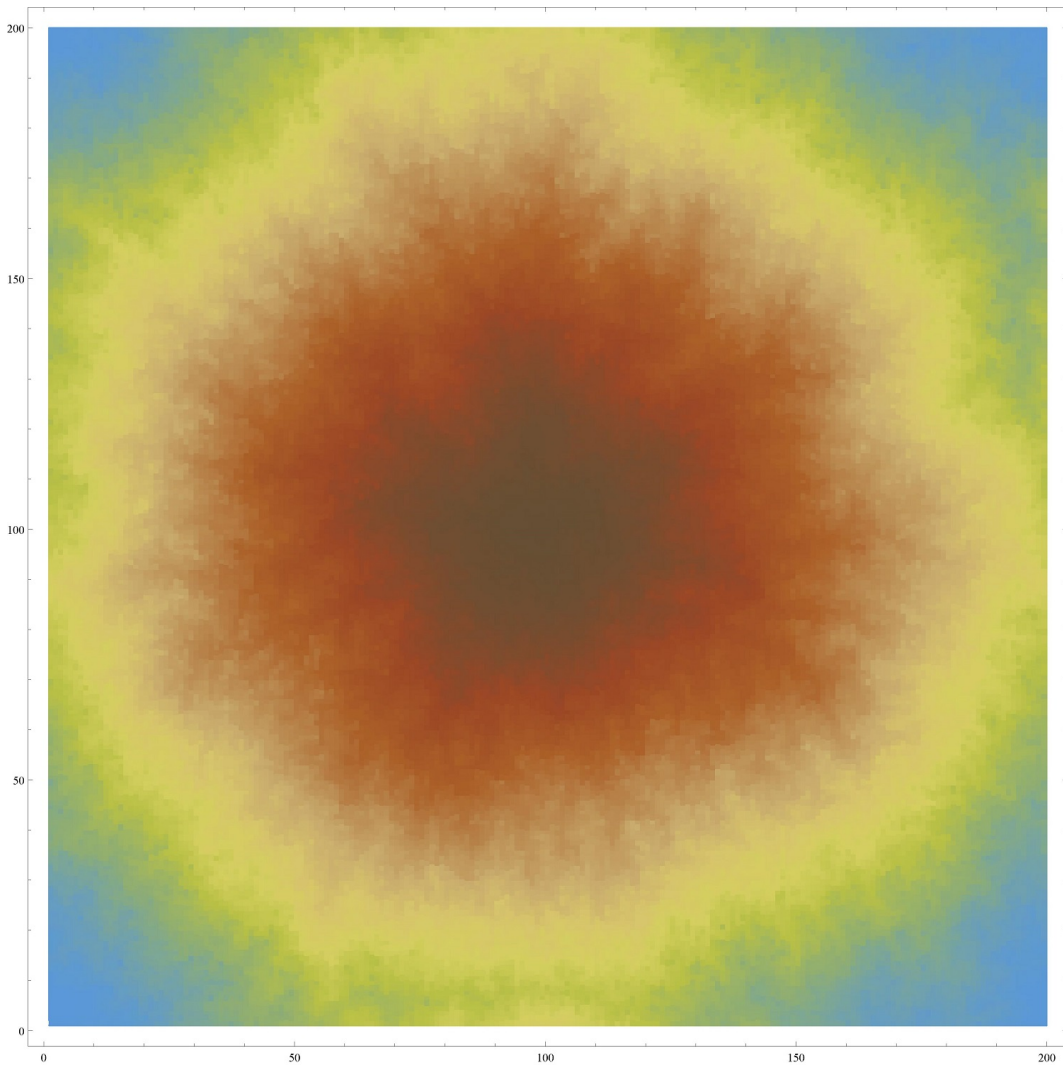


FIGURE 1. Eden model on the discrete torus with edge length 200. The original cluster was a single point at  $(100, 100)$ . The order of point attachment to the cluster is indicated by the warmth of the color.

In other words, if  $R(3)$  is chosen uniformly from all vertices adjacent to  $R(4)$  and this is followed by any choice of  $R(2)$  and  $R(1)$  consistent with the condition that  $R(4) = b$  is the only peak of  $L$  then the resulting labeling is not uniformly random (conditional on  $\{R(4) = b\}$ ).

### 3. PRELIMINARIES

For an integer  $N > 0$ , let  $[N] = \{1, 2, \dots, N\}$ . By an abuse of notation, we will use  $|\cdot|$  to denote the absolute value of a real number and cardinality of a finite set.

We will usually denote graphs by  $\mathcal{G}$ , their vertex sets by  $\mathcal{V}$  and edge sets by  $\mathcal{E}$ . We will indicate adjacency of vertices  $x$  and  $y$  by  $x \leftrightarrow y$ .

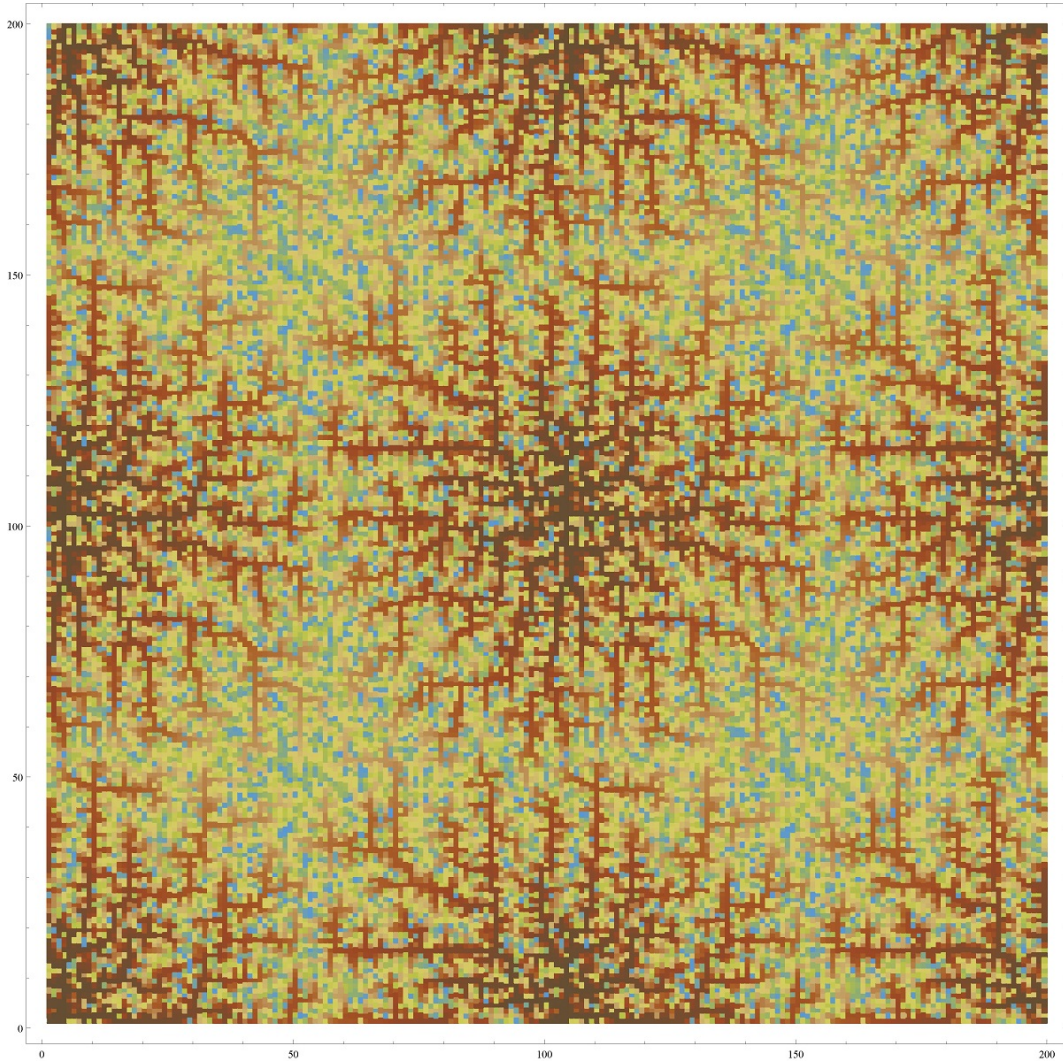


FIGURE 2. Random labeling of the discrete torus with edge length 100. The labeling is conditioned on having exactly one peak. Level lines are indicated by the warmth of the color. The figure consists of four copies of the same simulation, placed next to each other, to help see the main geometric features.

Suppose that  $1 < |\mathcal{V}| = N < \infty$ . We will call a function  $L : \mathcal{V} \rightarrow [N]$  a labeling if it is a bijection.

We will say that a vertex  $x \in \mathcal{V}$  is a peak (of  $L$ ) if and only if it is a local maximum of  $L$ , i.e.,  $L(x) > L(y)$  for all  $y \leftrightarrow x$ . The event that the number of peaks of a random labeling of graph  $\mathcal{G}$  is equal to  $k$  will be denoted  $\mathcal{N}(\mathcal{G}, k)$ .

We will use  $\mathbb{P}$  to denote the distribution of a random (uniform, unconditioned) labeling. The symbol  $\mathbb{P}_1$  ( $\mathbb{P}_2$ ) will stand for the distribution  $\mathbb{P}$  conditioned on existence of exactly one peak (exactly two peaks). If an argument involves multiple graphs, we may indicate the graph in the notation by writing, for example,  $\mathbb{P}_{\mathcal{G}}$  or  $\mathbb{P}_{1,\mathcal{G}}$ .

Recall the Stirling formula for  $n \geq 1$ ,

$$(3.1) \quad \log n! = n \log n - n + O(\log n).$$

This obviously implies that,

$$(3.2) \quad \log n! \leq n \log n + O(\log n).$$

A more accurate version of the Stirling formula says that, for  $n \geq 1$ ,

$$(3.3) \quad (n + 1/2) \log n - n + \log \sqrt{2\pi} \leq \log n! \leq (n + 1/2) \log n - n + 1.$$

For every  $k = 0, 1, \dots, n$ , we have  $\binom{n}{k} \leq \sum_{j=0}^n \binom{n}{j} = 2^n$ , so

$$(3.4) \quad \log \binom{n}{k} \leq n \log 2.$$

#### 4. ROUGHNESS OF LEVEL BOUNDARIES

Recall notation from Section 2. We will show that boundaries of clusters  $C_k$  (“level sets”) for random labelings of two-dimensional tori, conditioned on having a single peak, have dimension 2, in a suitable asymptotic sense.

**Remark 4.1.** Recall the Eden model, a sequence of randomly growing clusters  $\tilde{C}_k$ , from Section 2. The model was introduced in [Ede61]. The site version of the Eden model is identical to the site version of the “first passage percolation” model introduced in [HW65]. A shape theorem was proved in [CD81] for the edge version of the model; it says that when the size of the cluster goes to infinity, the rescaled cluster converges to a convex set  $B$ . For a connection between the Eden model and the KPZ equation, see, e.g., [Qua12].

Theorem 3.1 of [Ale97] (see also a stronger version in [DK14, Prop. 3.1]) implies that with probability 1, for all large  $n$ , the boundary of the set  $\tilde{C}_n$  lies in the “annulus” of width  $c_1 n^{1/4} \log n$  containing  $n^{1/2} \partial B$  or, more precisely, a neighborhood of width  $c_1 n^{1/4} \log n$  of the boundary of  $n^{1/2} B$ . This implies that the cardinality of the boundary of  $\tilde{C}_n$  is bounded by the volume of the annulus, that is,  $c_2 n^{3/4} \log n$ .

The following conjecture was communicated to us by Michael Damron. For some constant  $c_3$ , with probability one, there exist infinitely many  $n$  such that the boundary of  $\tilde{C}_n$  contains at most  $c_3 n^{1/2}$  vertices.

Even the weaker upper bound,  $c_2 n^{3/4} \log n$ , on the size of cluster boundaries in the Eden model and Theorem 4.2 show that the clusters  $C_k$  in the random labeling model have much different character from those in the Eden model. The comparison of Figs. 1 and 2 seems to support this claim.

Recall that  $\mathcal{G}_{n,n}$  denotes the Cartesian product of path (linear) graphs, each with  $n$  vertices. In other words,  $\mathcal{G}_{n,n}$  is a discrete square with edge length  $n$ . The set of vertices of  $\mathcal{G}_{n,n}$  is denoted  $\mathcal{V}_{n,n}$ .

Recall clusters  $C_k$  defined in Section 2. Let  $\partial C_k$  denote the set of vertices outside  $C_k$  that are adjacent to a vertex in  $C_k$ .

**Theorem 4.2.** *Consider a random labeling of  $\mathcal{G}_{n,n}$  conditioned on having a single peak. Let  $f(n) = (\log n)^{-1/2} \log \log n$ . Let  $F$  be the event that the number of  $k \in [n^2]$  such that  $|\partial C_k| \leq n^{2-f(n)}$  is greater than or equal to  $f(n)n^2$ . For large  $n$ ,*

$$\mathbb{P}_1(F) \leq \exp(-(1/2)n^2(\log \log n)^2).$$

*Proof. Step 1.* In this step, we will prove that for any  $1 \leq m_1, m_2 \leq n$ , the number of labelings of  $\mathcal{G}_{n,n}$  which have only one peak at  $(m_1, m_2)$  is greater than

$$(4.1) \quad \exp(2n^2 \log n + o(n^2 \log n)).$$

Let  $\ell = \ell(n) = \lfloor \log n \rfloor$ . Let  $\mathcal{H}_n$  be the graph whose vertex set is a subset of  $\mathcal{V}_{n,n}$  and contains all points with coordinates  $(k_1, k_2)$  satisfying at least one of the following four conditions: (i)  $k_1 = m_1$ , (ii)  $k_2 = m_2$ , (iii)  $k_1 = j\ell$  for some integer  $j$ , (iv)  $k_2 = j\ell$  for some integer  $j$ . Two vertices of  $\mathcal{H}_n$  are connected by an edge if and only if they are adjacent in  $\mathcal{G}_{n,n}$ . The graph  $\mathcal{H}_n$  has the shape of a lattice with the edge length  $\ell$  and a pair of extra lines passing through  $(m_1, m_2)$ .

Let  $N$  be the number of vertices in  $\mathcal{H}_n$ . We have  $N \leq 3n^2/\log n$  for large  $n$ . Let  $\mathcal{V}_n^{\mathcal{H}}$  denote the set of vertices of  $\mathcal{H}_n$ .

We will define a family of deterministic labelings  $L$  of  $\mathcal{G}_{n,n}$ . We will first label  $\mathcal{V}_n^{\mathcal{H}}$ . It is easy to see that one can label  $\mathcal{V}_n^{\mathcal{H}}$  in such a way that  $L(m_1, m_2) = n^2$ ,  $(m_1, m_2)$  is the only peak of  $L$ , and  $L(\mathcal{V}_n^{\mathcal{H}}) = \{n^2, n^2 - 1, \dots, n^2 - N + 1\}$ . Fix any such labeling of  $\mathcal{V}_n^{\mathcal{H}}$  and consider labelings  $L$  of  $\mathcal{V}_{n,n}$  which agree with this labeling on  $\mathcal{V}_n^{\mathcal{H}}$ .

Consider labelings with no restriction on peaks outside  $\mathcal{V}_n^{\mathcal{H}}$ . There are  $n^2 - N$  numbers in  $[n^2]$  not used for labeling  $\mathcal{V}_n^{\mathcal{H}}$  so there are  $(n^2 - N)!$  labelings of  $\mathcal{G}_{n,n}$  which extend a given labeling of  $\mathcal{V}_n^{\mathcal{H}}$ . Let this family be called  $\mathcal{A}$ .

Consider a maximal horizontal line segment  $\Gamma \subset \mathcal{V}_{n,n} \setminus \mathcal{V}_n^{\mathcal{H}}$ . The length of  $\Gamma$  is at most  $\ell - 1$ . At least one of the permutations of  $L(\Gamma)$  is monotone with the largest number adjacent to  $\mathcal{V}_n^{\mathcal{H}}$ . So there are at least  $(n^2 - N)!/(\ell - 1)!$  labelings in  $\mathcal{A}$  with no peaks on  $\Gamma$ . Let  $\Gamma_1$  be a maximal horizontal line segment in  $\mathcal{V}_{n,n} \setminus \mathcal{V}_n^{\mathcal{H}}$ , disjoint from  $\Gamma$ . We can apply the same argument using this time permutations of  $L(\Gamma_1)$  to see that the number of labelings in  $\mathcal{A}$  with no peaks on  $\Gamma \cup \Gamma_1$  is greater than or equal to  $(n^2 - N)!((\ell - 1)!)^{-2}$ . The set  $\mathcal{V}_{n,n} \setminus \mathcal{V}_n^{\mathcal{H}}$  can be partitioned into at most  $n(n+1)/\ell \leq (n^2/\log n)(1+1/n)(1+2/\log n)$  disjoint maximal horizontal line segments of length at most  $\ell$ , for large  $n$ . So the number of labelings in  $\mathcal{A}$  with no peaks is greater than or equal to  $(n^2 - N)!((\ell - 1)!)^{-(n^2/\log n)(1+1/n)(1+2/\log n)}$ . For large  $n$ , the logarithm of this number is greater than or equal to, using (3.3),

$$\begin{aligned} & (n^2 - N + 1/2) \log(n^2 - N) - (n^2 - N) + c_1 \\ & - (n^2/\log n)(1 + 1/n)(1 + 2/\log n)(\ell - 1 + 1/2) \log(\ell - 1) \\ & \geq (n^2 - 3n^2/\log n + 1/2) \log(n^2 - 3n^2/\log n) - n^2 \\ & - (n^2/\log n)(1 + 1/n)(1 + 2/\log n) \log n \log \log n \\ & \geq 2n^2(1 - 3/\log n) \log n + n^2(1 - 3/\log n) \log(1 - 3/\log n) - n^2 \\ & - n^2(1 + 1/n)(1 + 2/\log n) \log \log n \\ & \geq 2n^2 \log n + O(n^2 \log \log n). \end{aligned}$$



This proves (4.1).

*Step 2.* Suppose that  $B_1, B_2, \dots, B_{n^2}$  is a sequence of sets in  $\mathcal{V}_{n,n}$ , starting with a singleton  $B_1$  and such that  $|B_k| = |B_{k-1}| + 1$  for all  $k$ . Let  $\partial B_k$  denote the set of vertices outside  $B_k$  that are adjacent to a vertex in  $B_k$ . We also require that  $B_{k+1} \setminus B_k \subset \partial B_k$  for all  $k$ . Recall that  $f(n) = (\log n)^{-1/2} \log \log n$ . Let  $M$  be the number of sequences  $(B_k)_{k \in [n^2]}$  with the property that the number  $R$  of  $k \in [n^2]$  such that  $|\partial B_k| \leq n^{2-f(n)}$  is greater than or equal to  $f(n)n^2$ .

The number of possible values of  $R$  is bounded by  $n^2$ . Given  $R \geq f(n)n^2$ , the number of possible sequences of  $k$ 's such that  $|\partial B_k| \leq n^{2-f(n)}$  is bounded by  $\binom{n^2}{R}$ . Given a specific set  $\Lambda$  of  $k$ 's in  $[n^2]$  with cardinality greater than or equal to  $f(n)n^2$ , the number of sequences  $(B_k)_{k \in [n^2]}$  with the property that  $|\partial B_k| \leq n^{2-f(n)}$  for all  $k \in \Lambda$  is less than or equal to

$$(n^{2-f(n)})^{f(n)n^2} (n^2)^{(1-f(n))n^2+1},$$

so

$$M \leq n^2 \binom{n^2}{R} (n^{2-f(n)})^{f(n)n^2} (n^2)^{(1-f(n))n^2+1}.$$

By (3.4), for some  $c$  and all  $n$  and  $R$ ,  $\log \binom{n^2}{R} \leq cn^2$ . Hence,

$$\begin{aligned} \log M &\leq 2 \log n + cn^2 + f(n)n^2(2 - f(n)) \log n + 2((1 - f(n))n^2 + 1) \log n \\ &= cn^2 - n^2 f^2(n) \log n + 2n^2 \log n + 4 \log n \\ &= cn^2 - n^2 (\log n)^{-1} (\log \log n)^2 \log n + 2n^2 \log n + 4 \log n \\ &= cn^2 - n^2 (\log \log n)^2 + 2n^2 \log n + 4 \log n. \end{aligned}$$

The probability of  $F$  is less than or equal to  $M$  divided by the number in (4.1), so for some  $c$  this probability is bounded above by

$$\begin{aligned} &\frac{\exp(cn^2 - n^2 (\log \log n)^2 + 2n^2 \log n + 4 \log n)}{\exp(2n^2 \log n - cn^2 \log \log n)} \\ &= \exp(-n^2 (\log \log n)^2 + cn^2 + 4 \log n + cn^2 \log \log n) \\ &\leq \exp(-(1/2)n^2 (\log \log n)^2), \end{aligned}$$

for all large  $n$ . □

## 5. LADDER GRAPHS

The results of a simulation presented in Fig. 2 suggest that gradient lines for the random labeling contain long straight stretches. We cannot prove this feature of random labelings of tori but we will prove a similar result for “ladder” graphs. Theorem 5.1 will show that the longest of all gradient lines of a random labeling with a single peak of a ladder graph mostly consists of straight stretches, with the exception of a logarithmically small percentage of steps.

A generalized ladder graph  $\mathcal{G}_{m,n}$  is the Cartesian product of path (linear) graphs with  $m$  and  $n$  vertices. In other words,  $\mathcal{G}_{m,n}$  is a discrete rectangle with edge lengths  $m$  and

$n$ . We will consider  $m$  a fixed parameter and we will state an asymptotic theorem when  $n$  goes to infinity.

We will identify the set of vertices  $\mathcal{V}_{m,n}$  of the graph  $\mathcal{G}_{m,n}$  with points of  $\mathbb{Z}^2$ , so that we can refer to them using Cartesian coordinates  $(j, k)$ ,  $1 \leq j \leq n$ ,  $1 \leq k \leq m$ . Let  $\mathcal{B}$  be the family of all (deterministic) labelings of  $\mathcal{G}_{m,n}$  with only one peak at  $(1, 1)$ . Let  $\Lambda$  be the set of all “continuous” non-self intersecting paths taking values in  $\mathcal{V}_{m,n}$ , i.e.,  $\lambda \in \Lambda$  if and only if there is some  $k$  such that  $\lambda : [k] \rightarrow \mathcal{V}_{m,n}$ ,  $\lambda(j) \leftrightarrow \lambda(j+1)$  for  $j \in [k-1]$ , and  $\lambda(i) \neq \lambda(j)$  for  $i \neq j$ . We will write  $|\lambda| = k$ .

We will argue that for any  $L \in \mathcal{B}$  there exists  $\lambda_L = (\lambda_L^1, \lambda_L^2) \in \Lambda$  such that  $\lambda_L(1) = (1, 1)$ ,  $\lambda_L^1(|\lambda_L|) = n$ , and  $j \rightarrow L(\lambda_L(j))$  is a decreasing function. Fix any ordering of  $\mathcal{V}_{m,n}$ . We will construct a tree. Let  $v_1 = (1, 1)$  be the root of the tree. Suppose that vertices  $\{v_1, v_2, \dots, v_j\}$  of the tree have been chosen. Consider the following cases. (a) The set  $\mathcal{V}_{m,n} \setminus \{v_1, v_2, \dots, v_j\}$  contains a vertex  $w$  such that  $w$  is adjacent to a vertex in  $\{v_1, v_2, \dots, v_j\}$ , say,  $v_i$ , and  $L(w) < L(v_i)$ . Then we take the largest  $w$  with this property and add this  $w$  to  $\{v_1, v_2, \dots, v_j\}$  as a new element  $v_{j+1}$ . We also choose the largest  $v_i \in \{v_1, v_2, \dots, v_j\}$  with the property that  $L(v_{j+1}) < L(v_i)$  and add an edge to the tree between  $v_{j+1}$  and  $v_i$ . (b) Suppose that the set  $\mathcal{V}_{m,n} \setminus \{v_1, v_2, \dots, v_j\}$  does not contain a vertex  $w$  such that  $w$  is adjacent to a vertex in  $\{v_1, v_2, \dots, v_j\}$  and has the property that  $L(w) < L(v_i)$  for some  $v_i$ . If  $\{v_1, v_2, \dots, v_j\}$  contains a vertex of the form  $(n, k)$  then the unique path within the tree from  $v_1 = (1, 1)$  to  $(n, k)$  satisfies all conditions that  $\lambda_L$  is supposed to satisfy. In the opposite case,  $L$  restricted to the set  $\mathcal{V}_{m,n} \setminus \{v_1, v_2, \dots, v_j\}$ , must attain the maximum, say, at  $x$ . If  $x$  is adjacent to any vertex  $y \in \{v_1, v_2, \dots, v_j\}$  then  $L(x) > L(y)$  for every such  $y$  because of the assumption made at the beginning of case (b). This implies that  $x$  is a peak of  $L$  in the whole graph  $\mathcal{G}_{m,n}$ . Since this contradicts the assumption that  $L$  has only one peak at  $(1, 1)$ , the proof of existence of  $\lambda_L$  is complete.

If there is more than one labeling  $\lambda_L$  satisfying the properties stated above then we let  $\lambda_L$  be the labeling that, in addition, is minimal in some fixed arbitrary ordering of  $\Lambda$ .

**Theorem 5.1.** *Let  $L$  be a random labeling of  $\mathcal{V}_{m,n}$ , chosen uniformly from  $\mathcal{B}$  and let  $\mathbb{P}_{\mathcal{B}}$  denote the corresponding probability. For every  $m \geq 3$  there exists  $n_1$  such that for  $n \geq n_1$ ,*

$$\mathbb{P}_{\mathcal{B}} \left( |\lambda_L| \geq n \left( 1 + \frac{\log \log n}{\log n} \right) \right) \leq \exp(-(1/2)n \log \log n).$$

*Proof.* First we will estimate the total number of deterministic labelings of  $\mathcal{V}_{m,n}$  with a single peak at  $(1, 1)$ .

Let  $\mathcal{A}$  be the family of all deterministic labelings  $L$  of  $\mathcal{V}_{m,n}$  such that

$$L(1, 1) = nm, \quad L(2, 1) = nm - 1, \dots, \quad L(n, 1) = nm - n + 1.$$

Note that labelings in the family  $\mathcal{A}$  may have multiple peaks. There are  $(n(m-1))!$  labelings in  $\mathcal{A}$ .

Let  $\Gamma_k = \{(k, 2), (k, 3), \dots, (k, m)\}$  for  $k \in [n]$ . There exists a unique permutation of  $L(k, 2), L(k, 3), \dots, L(k, m)$  which is decreasing so the number of labelings in  $\mathcal{A}$  which do not have peaks on  $\Gamma_1$  is at least  $(n(m-1))!/(m-1)!$ . For the same reason, the number



of labelings in  $\mathcal{A}$  which do not have peaks on  $\Gamma_1 \cup \Gamma_2$  is at least  $(n(m-1))!/((m-1)!)^2$ . Extending the argument to all  $\Gamma_k$ 's, we conclude that the number of labelings in  $\mathcal{A}$  which do not have a peak on any  $\Gamma_k$  is at least  $(n(m-1))!/((m-1)!)^n$ . This implies that the number of labelings with a single peak at  $(1, 1)$  is bounded below by

$$(5.1) \quad (n(m-1))!/((m-1)!)^n.$$

Let  $f(n) = \log \log n / \log n$ . Consider any integer  $j$  such that  $n(1 + f(n)) \leq j \leq nm$ . We will count the number of paths  $\lambda \in \Lambda$  such that  $|\lambda| = j$  and  $\lambda = \lambda_L$  for some  $L \in \mathcal{B}$ . We must have  $L(1, 1) = nm$  and  $\lambda(1) = (1, 1)$ . There are at most  $\binom{nm}{j-1}$  choices for the values of  $L(\lambda(k))$ ,  $2 \leq k \leq j$ . Once these values are chosen, we arrange them along a “continuous” path by choosing inductively a vertex  $\lambda(k+1)$ , given  $\lambda(1), \dots, \lambda(k)$ . Going from the top value monotonically to the lowest value, there are at most 3 choices at every step so the number of  $\lambda$  such that  $|\lambda| = j$  and  $\lambda = \lambda_L$  for some  $L \in \mathcal{B}$  is bounded by  $\binom{nm}{j-1} 3^{j-1}$ .

If  $|\lambda| = j$  then there are  $nm - j$  vertices outside the range of  $\lambda$ . Hence, the number of labelings  $L \in \mathcal{B}$  such that  $\lambda_L$  is equal to a fixed  $\lambda$  with  $|\lambda| = j$  and the values of  $L(\lambda(k))$  are fixed for  $1 \leq k \leq j$ , is bounded by  $(nm - j)!$ . Let  $M(m, n, j)$  be the number of labelings  $L \in \mathcal{B}$  such that  $|\lambda_L| = j$ . We have

$$M(m, n, j) \leq \binom{nm}{j-1} 3^{j-1} (nm - j)! \leq \binom{nm}{j-1} 3^{j-1} (nm - j + 1)! = \frac{(nm)!}{(j-1)!} 3^{j-1}.$$

For  $n \geq 4$ , the expression on the right hand side, considered to be a function of  $j$  on the interval  $n(1 + f(n)) \leq j \leq nm$ , is maximized by  $j = \lceil n(1 + f(n)) \rceil$ . Let  $n_1 = \lceil n(1 + f(n)) \rceil - 1$ . We obtain the following bound,

$$M(m, n, j) \leq (nm)! 3^{n_1} / n_1!.$$

Next we sum over  $j$  such that  $n(1 + f(n)) \leq j \leq nm$ . The number of labelings  $L \in \mathcal{B}$  such that  $|\lambda_L| \geq n(1 + f(n))$  is bounded by

$$(5.2) \quad n(m - f(n))(nm)! 3^{n_1} / n_1!.$$

The probability that  $|\lambda_L| \geq n(1 + f(n))$  for a labeling chosen uniformly from  $\mathcal{B}$  is bounded above by the ratio of the numbers in (5.2) and (5.1). We use the Stirling formula (3.1) and (3.2) to see that for all sufficiently large  $n$ ,

$$\begin{aligned} & \mathbb{P}_{\mathcal{B}}(|\lambda_L| \geq n(1 + f(n))) \\ & \leq \frac{n(m - f(n))(nm)! 3^{n_1} / n_1!}{(n(m-1))!/((m-1)!)^n} \\ & = \frac{n(m - f(n))(nm)! 3^{n_1} ((m-1)!)^n}{(n(m-1))! n_1!} \\ & \leq \exp \left( \log n + \log(m - f(n)) + nm \log(nm) + O(\log(nm)) + n(1 + f(n)) \log 3 \right. \\ & \quad \left. + n(m-1) \log(m-1) + O(n \log(m-1)) \right. \\ & \quad \left. - n(m-1) \log(n(m-1)) + n(m-1) + O(\log(n(m-1))) \right. \\ & \quad \left. - n(1 + f(n)) \log(n(1 + f(n)) - 1) + n(1 + f(n)) + O(\log(n(1 + f(n)))) \right) \end{aligned}$$

$$\begin{aligned}
&\leq \exp \left( nm \log(nm) - n(m-1) \log n - n(1+f(n)) \log(n(1+f(n)) - 1) + O(n) \right) \\
&= \exp \left( nm \log n + nm \log m - n(m-1) \log n \right. \\
&\quad \left. - n \log n - nf(n) \log n - n \log(1+f(n) - 1/n) \right. \\
&\quad \left. - nf(n) \log(1+f(n) - 1/n) + O(n) \right) \\
&= \exp \left( -nf(n) \log n - n \log(1+f(n) - 1/n) - nf(n) \log(1+f(n) - 1/n) + O(n) \right) \\
&= \exp \left( -nf(n) \log n + O(n) \right) \\
&= \exp \left( -n(\log \log n / \log n) \log n + O(n) \right) \\
&\leq \exp(-(1/2)n \log \log n).
\end{aligned}$$

□

## 6. RANDOM LABELINGS OF TREES

It often happens that trees are more tractable than general graphs. This is indeed the case when we consider random labelings conditioned on a small number of peaks. We will start this section with some preliminary results.

**Definition 6.1.** (i) Suppose that a graph  $\mathcal{G}$  is a tree with  $N$  vertices. We will say that  $x$  is a centroid of  $\mathcal{G}$  if each subtree of  $\mathcal{G}$  which does not contain  $x$  has at most  $N/2$  vertices.

(ii) Suppose that  $\mathcal{G}$  is a tree and  $x$  is one of its vertices. Define a partial order “ $\leq$ ” on the set of vertices  $\mathcal{V}$  by declaring that for  $y, z \in \mathcal{V}$ , we have  $y \leq z$  if and only if  $y$  lies on the unique path joining  $x$  and  $z$ . If  $y \leq z$  then we will say that  $y$  is an ancestor of  $z$  and  $z$  is a descendant of  $y$ . We will write  $D_y^x$  to denote the family of all descendants of  $y$ , including  $y$ . We will also use the notation  $\mathbf{n}_x(y) = |D_y^x|$ . If needed, we will indicate the dependence on the graph by adding a superscript, for example,  $\mathbf{n}_x^{\mathcal{G}}(y)$ .

**Remark 6.2.** The following results can be found in [BH90, Thm. 2.3]. (i) A finite tree graph has at least one and at most two centroids. (ii) If it has two centroids then they are adjacent.

Part (i) cannot be strengthened to say that all trees have only one centroid. A linear graph with  $N$  vertices has one or two centroids depending on the parity of  $N$ .

**Proposition 6.3.** *Suppose that  $L$  is a random labeling of a tree  $\mathcal{G}$  and let  $K$  denote the location of the highest label. The function  $x \rightarrow \mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K = x\})$  attains the maximum at all centroids of  $\mathcal{G}$  and only at the centroids.*

*Proof.* Consider any  $x \in \mathcal{V}$ . Recall the notation  $D_y^x$  and  $\mathbf{n}_x(y)$  from Definition 6.1 (ii). Let

$$F_v = \left\{ L(v) \geq \max_{z \in D_v^x} L(z) \right\}.$$

Note that

$$\mathcal{N}(\mathcal{G}, 1) \cap \{K = x\} = \bigcap_{v \in \mathcal{V}} F_v.$$

Let  $\mathcal{V}_0 = \mathcal{V}_0^+ = \{x\}$ ,  $\mathcal{V}_k = \{v \in \mathcal{V} \setminus \mathcal{V}_{k-1}^+ : \exists z \in \mathcal{V}_{k-1}^+ \text{ such that } v \leftrightarrow z\}$  and  $\mathcal{V}_k^+ = \mathcal{V}_k \cup \mathcal{V}_{k-1}^+$  for  $k \geq 1$ . Let  $N = |\mathcal{V}|$  and note that  $\mathbb{P}(F_x) = 1/N = 1/\mathbf{n}_x(x)$ . It is easy to see that

$$\mathbb{P}\left(\bigcap_{v \leftrightarrow x} F_v \mid F_x\right) = \frac{1}{\prod_{v \leftrightarrow x} \mathbf{n}_x(v)}.$$

In general, for  $k \geq 1$ ,

$$\mathbb{P}\left(\bigcap_{v \in \mathcal{V}_k} F_v \mid \bigcap_{v \in \mathcal{V}_{k-1}^+} F_v\right) = \frac{1}{\prod_{v \in \mathcal{V}_k} \mathbf{n}_x(v)}.$$

It follows that

$$\begin{aligned} (6.1) \quad \mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K = x\}) &= \mathbb{P}(F_x) \prod_{k \geq 1} \mathbb{P}\left(\bigcap_{v \in \mathcal{V}_k} F_v \mid \bigcap_{v \in \mathcal{V}_{k-1}^+} F_v\right) \\ &= \prod_{k \geq 0} \frac{1}{\prod_{v \in \mathcal{V}_k} \mathbf{n}_x(v)} = \frac{1}{\prod_{z \in \mathcal{V}} \mathbf{n}_x(z)}. \end{aligned}$$

We will now vary  $x$ . Consider two adjacent vertices  $x$  and  $y$ . Note that  $\mathbf{n}_y(v) = \mathbf{n}_x(v)$  for  $v \neq x, y$ . We have

$$\mathbf{n}_y(x) = \mathbf{n}_x(x) - \mathbf{n}_x(y), \quad \mathbf{n}_y(y) = N = \mathbf{n}_x(x).$$

Therefore,

$$(6.2) \quad \frac{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K = x\})}{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K = y\})} = \frac{\mathbf{n}_y(x)\mathbf{n}_y(y)}{\mathbf{n}_x(x)\mathbf{n}_x(y)} = \frac{\mathbf{n}_y(x)}{\mathbf{n}_x(y)} = \frac{\mathbf{n}_x(x) - \mathbf{n}_x(y)}{\mathbf{n}_x(y)}.$$

It follows that  $\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K = x\}) \geq \mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K = y\})$  iff  $\mathbf{n}_x(x) - \mathbf{n}_x(y) \geq \mathbf{n}_x(y)$  iff  $N = \mathbf{n}_x(x) \geq 2\mathbf{n}_x(y)$  iff  $\mathbf{n}_x(y) \leq N/2$ . This can be easily rephrased as the statement of the proposition.  $\square$

Consider a tree  $\mathcal{G}$  with the vertex set  $\mathcal{V}$ . Let  $\mathbb{V}$  be the family of all (unordered) pairs of subsets  $\{\mathcal{V}', \mathcal{V}''\}$  of the vertex set  $\mathcal{V}$  such that  $\mathcal{V}'$  and  $\mathcal{V}''$  are vertex sets of non-empty subtrees of  $\mathcal{G}$ ,  $\mathcal{V}' \cap \mathcal{V}'' = \emptyset$  and  $\mathcal{V}' \cup \mathcal{V}'' = \mathcal{V}$ . In other words,  $\mathbb{V}$  is the family of all partitions of  $\mathcal{G}$  into two subtrees that can be generated by removing an edge.

**Lemma 6.4.** *Consider a (deterministic) labeling  $L$  of a tree  $\mathcal{G}$  that has exactly two peaks  $y_1$  and  $y_2$ .*

*(i) There exist exactly two pairs  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}$  such that  $y_1$  is the only peak of  $L$  restricted to  $\mathcal{V}'$  and  $y_2$  is the only peak of  $L$  restricted to  $\mathcal{V}''$ .*

*(ii) If the distance between  $y_1$  and  $y_2$  is greater than 2 then there exist at least three pairs  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}$  such that  $y_1 \in \mathcal{V}'$  and  $y_2 \in \mathcal{V}''$ . Suppose that  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}$  is one of such pairs and it does not satisfy the condition stated in (i). Then either (a)  $L$  restricted*

to  $\mathcal{V}$  has exactly two peaks, one of them located at  $y_1$  and the other one adjacent to  $\mathcal{V}''$ , and  $L$  restricted to  $\mathcal{V}''$  has exactly one peak at  $y_2$ , or (b)  $L$  restricted to  $\mathcal{V}''$  has exactly two peaks, one of them located at  $y_2$  and the other one adjacent to  $\mathcal{V}'$ , and  $L$  restricted to  $\mathcal{V}'$  has exactly one peak at  $y_1$ .

*Proof.* (i) Since  $y_1$  and  $y_2$  are peaks, the distance between them must be equal to or greater than 2. For  $x, z \in \mathcal{V}$ , let  $[x, z]$  denote the geodesic between  $x$  and  $z$ , including both vertices. If there are exactly two peaks  $y_1$  and  $y_2$  then there exists a unique  $y_3 \in [y_1, y_2]$  such that  $y_1 \neq y_3 \neq y_2$  and the labeling is monotone on both  $[y_1, y_3]$  and  $[y_2, y_3]$ . Let  $y_4$  be the neighbor of  $y_3$  in  $[y_1, y_3]$  and let  $y_5$  be the neighbor of  $y_3$  in  $[y_2, y_3]$ . Note that  $y_4$  could be  $y_1$  or  $y_5$  could be  $y_2$ . Let  $\{\mathcal{V}', \mathcal{V}''\}$  be subtrees obtained by removing the edge between  $y_4$  and  $y_3$  from  $\mathcal{G}$ . Let  $\{\tilde{\mathcal{V}}', \tilde{\mathcal{V}}''\}$  be subtrees obtained by removing the edge between  $y_5$  and  $y_3$  from  $\mathcal{G}$ . It is easy to see that  $\{\mathcal{V}', \mathcal{V}''\}$  and  $\{\tilde{\mathcal{V}}', \tilde{\mathcal{V}}''\}$  are the only elements of  $\mathbb{V}$  such that  $y_1$  is the only peak in one of the subtrees in the pair and  $y_2$  is the only peak in the other subtree in the same pair.

(ii) If the distance between  $y_1$  and  $y_2$  is greater than 2 then there are at least three edges along the geodesic  $[y_1, y_2]$ . Removing an edge not adjacent to  $y_3$  will generate a pair  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}$  with the properties specified in part (ii).  $\square$

The next result is related to the conjecture stated in Section 1. It follows easily from the arguments used in the proof of [BBPS15, Thm. 4.9] that if a random labeling of a large linear graph is conditioned to have exactly two peaks then the peaks are likely to be at a distance about one half of the length of the graph. Going in the opposite direction, we will show that the ratio of the distance between the twin peaks and the tree diameter can be arbitrarily close to zero.

In the following proposition, if the random labeling has exactly two peaks, we will denote the location of the highest peak  $K_1$  and the location of the other one  $K_2$ .

**Proposition 6.5.** *Fix arbitrarily small  $p > 0$  and arbitrarily large  $m$ . There exists a tree with diameter greater than  $m$  such that  $\mathbb{P}_2(K_1 = x, K_2 = y)$  attains the unique maximum for some  $x$  and  $y$  at the distance 3. Moreover,  $\mathbb{P}_2(\text{dist}(K_1, K_2) \geq 8) < p$ .*

*Proof.* Let  $m > 100$ ,  $n > (m!)^2$ , and let the vertex set  $\mathcal{V}$  of the graph  $\mathcal{G}$  consist of points  $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n, z_1, z_2, \dots, z_m$ . The only pairs of vertices connected by edges are  $(z_k, z_{k+1})$  for  $k \in [m-1]$ ,  $(z_1, x_k)$  for  $k \in [n]$  and  $(z_4, y_k)$  for  $k \in [n]$ .

First, we will find a lower estimate for

$$\mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = z_1, K_2 = z_4\} \cup \{K_2 = z_1, K_1 = z_4\})).$$

Let  $\mathcal{G}'$  be the subgraph of  $\mathcal{G}$  consisting of vertices  $z_1, z_2, x_1, x_2, \dots, x_n$ . Let  $\mathcal{G}''$  be the subgraph of  $\mathcal{G}$  consisting of all the remaining vertices. Let  $K'_1$  be the highest peak of a random labeling of  $\mathcal{G}'$  and let  $K''_1$  be the highest peak of a random labeling of  $\mathcal{G}''$ . We have

$$\begin{aligned} & \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = z_1, K_2 = z_4\} \cup \{K_2 = z_1, K_1 = z_4\})) \\ &= \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}', 1) \cap \{K'_1 = z_1\} \cap \mathcal{N}(\mathcal{G}'', 1) \cap \{K''_1 = z_4\}) \\ &= \mathbb{P}_{\mathcal{G}'}(\mathcal{N}(\mathcal{G}', 1) \cap \{K'_1 = z_1\}) \mathbb{P}_{\mathcal{G}''}(\mathcal{N}(\mathcal{G}'', 1) \cap \{K''_1 = z_4\}). \end{aligned}$$

We have equality on the second line above, despite Lemma 6.4 (ii), because if  $z_1$  and  $z_4$  are peaks of a labeling of  $\mathcal{G}$  then the same labeling restricted to  $\mathcal{G}'$  cannot have a peak at  $z_2$  and the labeling restricted to  $\mathcal{G}''$  cannot have a peak at  $z_3$ .

By (6.1) applied to  $\mathcal{G}'$ ,

$$\mathbb{P}_{\mathcal{G}'}(\mathcal{N}(\mathcal{G}', 1) \cap \{K'_1 = z_1\}) = \frac{1}{n+2}.$$

The same formula applied to  $\mathcal{G}''$  yields

$$\begin{aligned} \mathbb{P}_{\mathcal{G}''}(\mathcal{N}(\mathcal{G}'', 1) \cap \{K''_1 = z_4\}) &= \frac{1}{(n+m-2)(m-4)(m-5)\dots 2 \cdot 1} \\ &= \frac{1}{(n+m-2)(m-4)!}. \end{aligned}$$

It follows that

$$\begin{aligned} (6.3) \quad \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = z_1, K_2 = z_4\} \cup \{K_2 = z_1, K_1 = z_4\})) \\ \geq \frac{1}{(n+2)(n+m-2)(m-4)!}. \end{aligned}$$

Next we will find an upper bound for  $\mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_1, K_2 = v_2\})$ , assuming that  $(v_1, v_2) \neq (z_1, z_4)$  and  $(v_1, v_2) \neq (z_4, z_1)$ . Let  $\mathbf{G}$  be the family of all pairs  $(\mathcal{G}', \mathcal{G}'')$  of disjoint subtrees of  $\mathcal{G}$ , with vertex sets  $\mathcal{V}'$  and  $\mathcal{V}''$ , such that  $\mathcal{V}' \cup \mathcal{V}'' = \mathcal{V}$ ,  $v_1 \in \mathcal{V}'$  and  $v_2 \in \mathcal{V}''$ . Let  $K'_1$  be the highest peak of a random labeling of  $\mathcal{G}'$  and let  $K'_2$  be the second highest peak of  $\mathcal{G}'$ ; similar notation will be used for  $\mathcal{G}''$ . We will use Lemma 6.4.

Consider a pair  $(\mathcal{G}', \mathcal{G}'') \in \mathbf{G}$ . Note that if  $z_1$  is in  $\mathcal{G}'$  then at least  $n-1$  of vertices  $x_1, x_2, \dots, x_n$  are also in  $\mathcal{G}'$ . If  $z_4$  is in  $\mathcal{G}'$  then at least  $n-1$  of vertices  $y_1, y_2, \dots, y_n$  are also in  $\mathcal{G}'$ . Analogous claims hold for  $z_1$  and  $\mathcal{G}''$ , and for  $z_4$  and  $\mathcal{G}''$ .

Suppose that  $z_1$  and  $z_4$  are in two different graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ . Suppose without loss of generality that  $z_1$  is in  $\mathcal{G}'$  and  $z_4$  is in  $\mathcal{G}''$ . Recall that  $(v_1, v_2) \neq (z_1, z_4)$ . Consider the case when  $v_1 \neq z_1$ . Note that  $\mathbf{n}_{v_1}^{\mathcal{G}'}(v_1) \geq n$  and  $\mathbf{n}_{v_1}^{\mathcal{G}'}(z_1) \geq n-1$  so, by (6.1),

$$\mathbb{P}_{\mathcal{G}'}(\mathcal{N}(\mathcal{G}', 1) \cap \{K'_1 = v_1\}) \leq \frac{1}{n(n-1)}.$$

Since  $\mathbf{n}_{v_2}^{\mathcal{G}''}(v_2) \geq n$  so, by (6.1),

$$\mathbb{P}_{\mathcal{G}''}(\mathcal{N}(\mathcal{G}'', 1) \cap \{K''_1 = v_2\}) \leq \frac{1}{n}.$$

A completely analogous argument shows that, if  $v_1 = z_1$  and  $v_2 \neq z_4$  then

$$\begin{aligned} (6.4) \quad \mathbb{P}_{\mathcal{G}'}(\mathcal{N}(\mathcal{G}', 1) \cap \{K'_1 = v_1\}) &\leq \frac{1}{n}, \\ \mathbb{P}_{\mathcal{G}''}(\mathcal{N}(\mathcal{G}'', 1) \cap \{K''_1 = v_2\}) &\leq \frac{1}{n(n-1)}. \end{aligned}$$

We conclude that if  $z_1 \in \mathcal{V}'$ ,  $z_4 \in \mathcal{V}''$  and  $(v_1, v_2) \neq (z_1, z_4)$  then,

$$(6.5) \quad \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}', 1) \cap \mathcal{N}(\mathcal{G}'', 1) \cap \{K'_1 = v_1, K''_1 = v_2\}) \leq \frac{1}{n^2(n-1)}.$$

Recall the possibilities listed in Lemma 6.4. We will estimate the probability of exactly two peaks in one of the subgraphs.

We are returning to the case when  $z_1 \in \mathcal{V}'$ ,  $z_4 \in \mathcal{V}''$ ,  $(v_1, v_2) \neq (z_1, z_4)$  and  $v_1 \neq z_1$ . Since  $\mathcal{V}'$  contains at least  $n$  vertices,  $\mathbb{P}_{\mathcal{G}'}(K'_1 = v_1) \leq 1/n$ . If the labeling has exactly two peaks then there are at least  $n-2$  vertices  $x'_1, \dots, x'_{n-2}$  among  $x_1, \dots, x_n$ , such that  $x'_k \in \mathcal{V}'$  and  $x'_k$  is not a peak, for  $k = 1, \dots, n-2$ . Therefore, the label of  $z_1$  must be larger than the labels of all  $x'_k$ 's. Note that  $n \leq |\mathcal{V}'| \leq n+3$ . Assuming that  $K'_1 = v_1$ , the probability that the label of  $z_1$  is larger than the labels of all  $x'_k$ 's is bounded by  $5/(n-1)$ . Hence,

$$\mathbb{P}_{\mathcal{G}'}(\mathcal{N}(\mathcal{G}', 2) \cap \{K'_1 = v_1\}) \leq \frac{5}{n(n-1)}.$$

The graph  $\mathcal{G}''$  has at least  $n$  vertices so  $\mathbb{P}_{\mathcal{G}''}(K''_1 = v_2) \leq 1/n$ . We see that if  $z_1 \in \mathcal{V}'$ ,  $z_4 \in \mathcal{V}''$ ,  $(v_1, v_2) \neq (z_1, z_4)$  and  $v_1 \neq z_1$  then,

$$(6.6) \quad \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}', 2) \cap \mathcal{N}(\mathcal{G}'', 1) \cap \{K'_1 = v_1, K''_1 = v_2\}) \leq \frac{5}{n^2(n-1)}.$$

The next case to consider is when  $z_1 \in \mathcal{V}'$ ,  $z_4 \in \mathcal{V}''$ ,  $(v_1, v_2) \neq (z_1, z_4)$  and  $v_2 \neq z_4$ . Since  $\mathcal{V}''$  contains at least  $n$  vertices,  $\mathbb{P}_{\mathcal{G}''}(K''_1 = v_2) \leq 1/n$ . If the labeling has exactly two peaks then there are at least  $n-2$  vertices  $y'_1, \dots, y'_{n-2}$  among  $y_1, \dots, y_n$ , such that  $y'_k \in \mathcal{V}''$  and  $y'_k$  is not a peak, for  $k = 1, \dots, n-2$ . Therefore, the label of  $z_4$  must be larger than the labels of all  $y'_k$ 's. Note that  $n \leq |\mathcal{V}''| \leq n+m$ . Assuming that  $K''_1 = v_2$ , the probability that the label of  $z_4$  is larger than the labels of all  $y'_k$ 's is bounded by  $(m+2)/(n-1)$ . Hence,

$$\mathbb{P}_{\mathcal{G}''}(\mathcal{N}(\mathcal{G}'', 2) \cap \{K''_1 = v_2\}) \leq \frac{m+2}{n(n-1)}.$$

The graph  $\mathcal{G}'$  has at least  $n$  vertices so  $\mathbb{P}_{\mathcal{G}'}(K'_1 = v_1) \leq 1/n$ . We see that if  $z_1 \in \mathcal{V}'$ ,  $z_4 \in \mathcal{V}''$ ,  $(v_1, v_2) \neq (z_1, z_4)$  and  $v_2 \neq z_4$  then,

$$(6.7) \quad \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}', 1) \cap \mathcal{N}(\mathcal{G}'', 2) \cap \{K'_1 = v_1, K''_1 = v_2\}) \leq \frac{m+2}{n^2(n-1)}.$$

Next suppose that  $z_1$  and  $z_4$  are in the same of two graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ . Without loss of generality, suppose that they are in  $\mathcal{G}'$ . It is easy to see that  $\mathbf{n}_{v_1}^{\mathcal{G}'}(v_1) \geq n$  and  $\mathbf{n}_{v_1}^{\mathcal{G}'}(x) \geq n$  for at least two other  $x$  in the set  $\{z_1, z_2, z_3, z_4\}$  so, by (6.1),

$$\mathbb{P}_{\mathcal{G}'}(K'_1 = v_1) \leq \frac{1}{n^3}.$$

The maximum of  $1/n^3$  and the bounds in (6.5), (6.6) and (6.7) is less than  $\frac{m+2}{n^2(n-1)}$  for  $m > 3$ . We obtain for any fixed  $(\mathcal{G}', \mathcal{G}'') \in \mathbf{G}$ ,

$$\mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap \{K'_1 = v_1, K''_1 = v_2\}) \leq \frac{m+2}{n^2(n-1)}.$$

It is easy to see that for any  $v_1$  and  $v_2$  which are not adjacent, the family  $\mathbf{G}$  has at most  $m$  elements. It follows that

$$\begin{aligned} \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_1, K_2 = v_2\}) &\leq \sum_{(\mathcal{G}', \mathcal{G}'') \in \mathbf{G}} \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap \{K'_1 = v_1, K''_1 = v_2\}) \\ &\leq m \frac{m+2}{n^2(n-1)} \leq \frac{m(m+2)}{n(n-1)(m!)^2}. \end{aligned}$$

Comparing this bound to (6.3), we conclude that  $\mathbb{P}_2(K_1 = x, K_2 = y)$  is maximized at  $(x, y) = (z_1, z_4)$  or  $(x, y) = (z_4, z_1)$ , for all large  $m$ .

We will now strengthen our estimates assuming that  $\text{dist}(v_1, v_2) \geq 8$ .

First suppose that  $z_1$  and  $z_4$  are in the same of two graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ . Without loss of generality, suppose that they are in  $\mathcal{G}'$ . Then  $\mathbf{n}_{v_1}^{\mathcal{G}'}(x) \geq n-1$  for all  $x$  in the set  $\{z_1, z_2, z_3, z_4\}$  (this is true whether  $v_1$  belongs to this set or not) so, by (6.1),

$$(6.8) \quad \mathbb{P}_{\mathcal{G}'}(K'_1 = v_1) \leq \frac{1}{(n-1)^4}.$$

Since  $z_1$  is in  $\mathcal{G}'$ , the condition  $\text{dist}(v_1, v_2) \geq 8$  implies that  $v_2 = z_k$  for some  $k \geq 7$ . It follows that there are at most  $6nm^2$  families  $(v_1, \mathcal{G}', v_2, \mathcal{G}'')$  such that  $\text{dist}(v_1, v_2) \geq 8$  and  $z_1$  and  $z_4$  are in the same of the two graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ .

Suppose that  $z_1$  and  $z_4$  are in two different graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ . Consider the case when  $z_1$  is in  $\mathcal{G}'$ . Recall that  $\text{dist}(v_1, v_2) \geq 8$  implies that  $v_2 = z_k$  for some  $k \geq 7$ . We have  $\mathbf{n}_{v_2}^{\mathcal{G}''}(v_2) \geq n$  and  $\mathbf{n}_{v_2}^{\mathcal{G}''}(z_k) \geq n$  for  $k = 4, 5, 6$ , so, by (6.1),

$$(6.9) \quad \mathbb{P}_{\mathcal{G}''}(\mathcal{N}(\mathcal{G}'', 1) \cap \{K''_1 = v_2\}) \leq \frac{1}{n^4}.$$

There are at most  $18nm$  families  $(v_1, \mathcal{G}', v_2, \mathcal{G}'')$  such that  $\text{dist}(v_1, v_2) \geq 8$  and  $z_1$  and  $z_4$  are in two different graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ .

We now analyze the case when there are exactly two peaks in  $\mathcal{G}''$ . Since  $\mathcal{V}''$  contains at least  $n$  vertices,  $\mathbb{P}_{\mathcal{G}''}(K''_1 = v_2) \leq 1/n$ . If the labeling has exactly two peaks then there are at least  $n-2$  vertices  $y'_1, \dots, y'_{n-2}$  among  $y_1, \dots, y_n$ , such that  $y'_k \in \mathcal{V}''$  and  $y'_k$  is not a peak, for  $k = 1, \dots, n-2$ . Therefore, the label of  $z_4$  must be larger than the labels of all  $y'_k$ 's. Recall that  $n \leq |\mathcal{V}''| \leq n+m$ . Assuming that  $K''_1 = v_2$ , the probability that the label of  $z_4$  is larger than the labels of all  $y'_k$ 's is bounded by  $(m+2)/(n-1)$ . Hence,

$$(6.10) \quad \mathbb{P}_{\mathcal{G}''}(\mathcal{N}(\mathcal{G}'', 2) \cap \{K''_1 = v_2\}) \leq \frac{m+2}{n(n-1)}.$$

The graph  $\mathcal{G}'$  has at least  $n$  vertices so  $\mathbb{P}_{\mathcal{G}'}(K'_1 = v_1) \leq 1/n$ . If  $v_1 = z_1$  then, using (6.10),

$$(6.11) \quad \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}', 1) \cap \mathcal{N}(\mathcal{G}'', 2) \cap \{K'_1 = v_1 = z_1, K''_1 = v_2\}) \leq \frac{m+2}{n^2(n-1)}.$$

There are at most  $6m$  families  $(v_1, \mathcal{G}', v_2, \mathcal{G}'')$  such that  $\text{dist}(v_1, v_2) \geq 8$ ,  $z_1$  and  $z_4$  are in two different graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ , and  $\mathcal{G}''$  and  $v_1 = z_1$  or  $v_2 = z_1$ .



If  $v_1 \neq z_1$  then  $\mathbf{n}_{v_1}^{\mathcal{G}'}(v_1) \geq n$  and  $\mathbf{n}_{v_1}^{\mathcal{G}'}(z_1) \geq n$ , so, by (6.1),

$$\mathbb{P}_{\mathcal{G}'}(\mathcal{N}(\mathcal{G}', 1) \cap \{K'_1 = v_1\}) \leq \frac{1}{n^2},$$

and, in view of (6.10),

$$(6.12) \quad \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}', 1) \cap \mathcal{N}(\mathcal{G}'', 2) \cap \{K'_1 = v_1, K''_1 = v_2\}) \leq \frac{m+2}{n^3(n-1)}.$$

There are at most  $18nm$  families  $(v_1, \mathcal{G}', v_2, \mathcal{G}'')$  such that  $\text{dist}(v_1, v_2) \geq 8$  and  $z_1$  and  $z_4$  are in two different graphs  $\mathcal{G}'$  and  $\mathcal{G}''$ .

We now appeal to Lemma 6.4 and combine (6.8), (6.9), (6.11) and (6.12) to conclude that, for large  $m$ ,

$$\begin{aligned} & \sum_{\text{dist}(v_1, v_2) \geq 8} \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_1, K_2 = v_2\}) \\ & \leq 6nm^2 \frac{1}{(n-1)^4} + 18nm \frac{1}{n^4} + 6m \frac{m+2}{n^2(n-1)} + 18nm \frac{m+2}{n^3(n-1)} \leq \frac{50m(m+2)}{(n-1)^3}. \end{aligned}$$

Therefore, in view of (6.3),

$$\begin{aligned} \mathbb{P}_2(\text{dist}(K_1, K_2) \geq 8) & \leq \frac{\sum_{\text{dist}(v_1, v_2) \geq 8} \mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_1, K_2 = v_2\})}{\mathbb{P}_{\mathcal{G}}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = z_1, K_2 = z_4\} \cup \{K_2 = z_1, K_1 = z_4\}))} \\ & \leq \frac{50m(m+2)}{(n-1)^3} (n+1)(n+m-2)(m-4)!. \end{aligned}$$

The last expression goes to 0 when  $n \rightarrow \infty$  so for any  $p > 0$ , we have  $\mathbb{P}_2(\text{dist}(K_1, K_2) \geq 8) < p$ , if  $n$  is sufficiently large.  $\square$

## 7. TWIN PEAKS ON REGULAR TREES

This section investigates twin peaks on  $(d+1)$ -regular trees of constant depth.

**Definition 7.1.** A finite rooted tree will be called a  $(d+1)$ -regular tree of depth  $k$  if the root, denoted  $v_*$ , has  $d+1$  children, every child has further  $d$  children and so on, continuing till  $k$  generations. That is, the distance of each leaf to the root is exactly  $k$ , and the degree of every non-leaf vertex is  $d+1$ .

Our main result in this section is not as complete as we wish it had been. Theorem 7.2 (i) requires the assumption that  $d \geq 10$ ; we believe that the assumption could be weakened.

We will briefly outline heuristic reasons why finding the location of the two peaks on a  $(d+1)$ -regular tree seems to be particularly challenging. If there are only two peaks on a tree graph, the graph can be divided into two subtrees such that the labeling restricted to each of the subtrees has only one or two peaks (see Lemma 6.4). If a  $(d+1)$ -regular tree, for  $d \geq 2$ , of constant depth is divided into two subtrees then it is easy to see that the centroid of the smaller subtree is adjacent to a vertex in the other subtree. Hence, in view of Proposition 6.3, one would expect the top peak in the smaller of the two subtrees to be close to the other subtree. This seems to contradict

the tendency for two peaks to be far away on “well-structured” graphs (see the remarks preceding Proposition 6.5).

**Theorem 7.2.** *Consider a random labeling of a  $(d+1)$ -regular tree  $\mathcal{G}$  of depth  $k$ . Let the locations of the highest and second highest peaks of the labeling be denoted  $K_1$  and  $K_2$ , respectively.*

(a) *If  $d \geq 10$  and  $k \geq 2$  then the function  $(y_1, y_2) \rightarrow \mathbb{P}_2(K_1 = y_1, K_2 = y_2)$  takes the maximum value only if  $y_1 = v_*$ .*

(b) *If  $d \geq 3$ ,  $k \geq 2$  and  $m \geq \sqrt{8k}$  then*

$$\mathbb{P}_2 \left( \min_{i=1,2} \text{dist}(K_i, v_*) \geq m \right) \leq 8 \exp(-(m/2) \log(d-1)).$$

(c) *For every  $d \geq 3$  there exists  $c_1 > 0$  such that for all  $k \geq 2$  and  $y_1, y_2 \in \mathcal{V}$  with  $\text{dist}(y_i, v_*) \geq 2$  for  $i = 1, 2$ , we have*

$$\mathbb{P}_2(K_1 = v_*, K_2 = y_1) \geq c_1 \mathbb{P}_2(K_1 = v_*, K_2 = y_2).$$

(d) *For every  $d \geq 3$  there exist  $c$  and  $c'$  such that for  $k \geq 2$  and  $n \geq 2\sqrt{8k}$ ,*

$$\begin{aligned} \mathbb{P}_2(|\text{dist}(K_1, K_2) - k| \geq n) \\ \leq 8 \exp(-(n/4) \log(d-1)) + c \exp(-(n/2) \log d) \\ \leq c' \exp(-(\sqrt{k}/64) \log d). \end{aligned}$$

*Proof.* (a) *Step 1.* First we will prove a general estimate similar to the “total probability formula.” Suppose that  $A_1, A_2, \dots, A_n$  satisfy the following condition,

$$A_i \cap A_j \cap A_m = \emptyset \quad \text{if } i \neq j \neq m \neq i.$$

These events need not be pairwise disjoint but no triplet has a nonempty intersection. Suppose that for some  $c_1 > 0$ , some events  $B$  and  $C$ , and all  $i$ ,

$$\mathbb{P}(B \mid A_i) \leq c_1 \mathbb{P}(C \mid A_i).$$

Assume that  $B \subset \bigcup_{i=1}^n A_i$ . We will show that

$$(7.1) \quad \mathbb{P}(B) \leq 2c_1 \mathbb{P}(C).$$

The proof is contained in the following calculation,

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P} \left( B \cap \bigcup_{i=1}^n A_i \right) = \mathbb{P} \left( \bigcup_{i=1}^n (B \cap A_i) \right) \leq \sum_{i=1}^n \mathbb{P}(B \cap A_i) = \sum_{i=1}^n \mathbb{P}(B \mid A_i) \mathbb{P}(A_i) \\ &\leq \sum_{i=1}^n c_1 \mathbb{P}(C \mid A_i) \mathbb{P}(A_i) = \sum_{i=1}^n c_1 \mathbb{P}(C \cap A_i) = c_1 \sum_{i=1}^n \mathbb{E} \mathbf{1}_{C \cap A_i} \\ &= c_1 \mathbb{E} \left( \sum_{i=1}^n \mathbf{1}_{C \cap A_i} \right) \leq c_1 \mathbb{E} (2 \cdot \mathbf{1}_C) = 2c_1 \mathbb{P}(C). \end{aligned}$$

*Step 2.* If we remove an edge (but retain all vertices) of  $\mathcal{G}$ , we obtain two subtrees. Consider a subtree  $\mathcal{G}'$  constructed in this way and let  $v_1$  be the vertex in  $\mathcal{G}'$  closest to the

root  $v_*$  in  $\mathcal{G}$ . If  $v_*$  is in  $\mathcal{G}'$  then  $v_1 = v_*$ . Consider a random labeling of  $\mathcal{G}'$  conditioned on having exactly one peak. Let  $K$  denote the position of the peak.

We create a branching structure in  $\mathcal{G}'$  by declaring that  $v_1$  is the ancestor of all vertices in  $\mathcal{G}'$ . Suppose that a vertex  $x_1 \in \mathcal{G}'$  is a parent of  $x_2 \in \mathcal{G}'$ . Recall the notation  $\mathbf{n}_x^{\mathcal{G}'}(y)$  from Definition 6.1. It is easy to see that  $\mathbf{n}_{x_2}^{\mathcal{G}'}(x_1) \geq (d-1)\mathbf{n}_{x_1}^{\mathcal{G}'}(x_2)$ . It follows from (6.2) that

$$(7.2) \quad \mathbb{P}_1(K = x_1) \geq (d-1) \mathbb{P}_1(K = x_2).$$

By induction, for any  $x \in \mathcal{G}'$  such that  $x \neq v_1$ ,

$$(7.3) \quad \mathbb{P}_1(K = v_1) \geq (d-1) \mathbb{P}_1(K = x).$$

We will also need sharper versions of the above estimates. Suppose that  $v_1 = v_*$ , i.e.,  $v_* \in \mathcal{G}'$ . Suppose that a vertex  $x_1 \in \mathcal{G}'$  is a parent of  $x_2 \in \mathcal{G}'$  and the distance from  $x_1$  to  $v_*$  is  $j$ . Then  $\mathbf{n}_{x_2}^{\mathcal{G}'}(x_1) \geq (d-1)^{j+1} \mathbf{n}_{x_1}^{\mathcal{G}'}(x_2)$ . It follows from (6.2) that

$$(7.4) \quad \mathbb{P}_1(K = x_1) \geq (d-1)^{j+1} \mathbb{P}_1(K = x_2).$$

By induction, for any  $x \in \mathcal{G}'$  such that  $\text{dist}(x, v_*) = m \geq 1$ ,

$$(7.5) \quad \mathbb{P}_1(K = v_*) \geq \mathbb{P}_1(K = x) \prod_{j=0}^{m-1} (d-1)^{j+1} = (d-1)^{m(m+1)/2} \mathbb{P}_1(K = x).$$

*Step 3.* Let  $\mathbb{V}$  be the family of all (unordered) pairs of subsets  $\{\mathcal{V}', \mathcal{V}''\}$  of the vertex set  $\mathcal{V}$  such that  $\mathcal{V}'$  and  $\mathcal{V}''$  are the vertex sets of non-empty subtrees,  $\mathcal{V}' \cap \mathcal{V}'' = \emptyset$  and  $\mathcal{V}' \cup \mathcal{V}'' = \mathcal{V}$ . In other words,  $\mathbb{V}$  is the family of all partitions of  $\mathcal{G}$  into two subtrees that can be generated by removing an edge. Graphs corresponding to  $\mathcal{V}'$  and  $\mathcal{V}''$  will be denoted  $\mathcal{G}'$  and  $\mathcal{G}''$ .

Consider the following conditions on vertices  $y_1, y_2 \in \mathcal{V}$ ,

(A1) The distance between  $y_1$  and  $y_2$  is larger than 2.

(A2) The distance between  $y_1$  and  $y_2$  is equal to 2 and the vertex between  $y_1$  and  $y_2$  is not  $v_*$ . We will denote this vertex by  $y_3$ .

In case (A1), let  $\mathbb{V}(y_1, y_2) \subset \mathbb{V}$  be the family of all  $\{\mathcal{V}', \mathcal{V}''\}$  such that  $y_k \in \mathcal{V}_k$  and  $y_k$  is not adjacent to a vertex in  $\mathcal{V}_{3-k}$ , for  $k = 1, 2$ .

In case (A2), we define the family  $\mathbb{V}(y_1, y_2)$  as a set containing only one pair  $\{\mathcal{V}', \mathcal{V}''\}$  constructed as follows. If we remove  $y_3$  then the graph is split into two subgraphs  $\mathcal{V}_1$  and  $\mathcal{V}_2$ , labeled so that  $y_k \in \mathcal{V}_k$  for  $k = 1, 2$ . Suppose that  $v_* \in \mathcal{V}_1$ . Then we let  $\mathcal{V}'' = \mathcal{V}_2 \cup \{y_3\}$  and  $\mathcal{V}' = \mathcal{V}_1$ . If  $v_* \in \mathcal{V}_2$  then we let  $\mathcal{V}' = \mathcal{V}_1 \cup \{y_3\}$  and  $\mathcal{V}'' = \mathcal{V}_2$ .

For  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}$  and a random labeling, let  $A(\mathcal{V}', \mathcal{V}'') = \mathcal{N}(\mathcal{G}', 1) \cap \mathcal{N}(\mathcal{G}'', 1)$ , i.e.,  $A(\mathcal{V}', \mathcal{V}'')$  is the event that the labeling restricted to  $\mathcal{V}_i$  has exactly one peak, for both  $i = 1$  and  $i = 2$ .

It follows from Lemma 6.4 that in both cases, (A1) and (A2),

$$(7.6) \quad \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_1, K_2 = y_2\} \subset \bigcup_{\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}(y_1, y_2)} A(\mathcal{V}', \mathcal{V}'').$$

*Step 4.* Let

$$B_1 = \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_1, K_2 = y_2\}, \quad B_2 = \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_2, K_2 = y_1\}.$$

Assume that (A1) or (A2) holds. Suppose that  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}(y_1, y_2)$ . Recall that  $y_1 \in \mathcal{V}'$ . We will show that if  $v_* \in \mathcal{V}'$  then

$$(7.7) \quad \mathbb{P}(B_1 \cap A(\mathcal{V}', \mathcal{V}'')) \geq \mathbb{P}(B_2 \cap A(\mathcal{V}', \mathcal{V}'')).$$

Let  $F$  be the event that the highest label is given to a vertex in  $\mathcal{V}'$ . Let  $K'_1$  be the location of the highest peak in  $\mathcal{G}'$  and let  $K''_1$  be the location of the highest peak in  $\mathcal{G}''$ .

Since  $v_* \in \mathcal{V}'$ , we have  $|\mathcal{V}'| > |\mathcal{V}''|$ , and thus,  $p_1 := \mathbb{P}(F) > 1/2$ . Note that

$$\begin{aligned} & \mathbb{P}(\mathcal{N}(\mathcal{G}', 1) \cap \mathcal{N}(\mathcal{G}'', 1) \cap \{K'_1 = y_1, K''_1 = y_2\} \mid F) \\ &= \mathbb{P}(\mathcal{N}(\mathcal{G}', 1) \cap \mathcal{N}(\mathcal{G}'', 1) \cap \{K'_1 = y_1, K''_1 = y_2\} \mid F^c), \end{aligned}$$

and call the common value of the two conditional probabilities  $p_2$ .

Since  $y_2$  is not adjacent to a vertex in  $\mathcal{V}'$ , the event  $\{B_1 \cap A(\mathcal{V}', \mathcal{V}'')\}$  is the same as the intersection of the events (i) the highest label is given to a vertex in  $\mathcal{V}'$ , (ii) there is a single peak at  $y_1$  in  $\mathcal{V}'$  and (iii) there is a single peak at  $y_2$  in  $\mathcal{V}''$ . Hence  $\mathbb{P}(B_1 \cap A(\mathcal{V}', \mathcal{V}'')) = p_1 p_2$ .

If the event  $\{B_2 \cap A(\mathcal{V}', \mathcal{V}'')\}$  holds then the intersection of the following events also holds: (i) the highest label is given to a vertex in  $\mathcal{V}''$ , (ii) there is a single peak at  $y_1$  in  $\mathcal{V}'$  and (iii) there is a single peak at  $y_2$  in  $\mathcal{V}''$  (but  $\{B_2 \cap A(\mathcal{V}', \mathcal{V}'')\}$  is not equal to the intersection of (i)-(iii)). This implies that  $\mathbb{P}(B_2 \cap A(\mathcal{V}', \mathcal{V}'')) \leq (1 - p_1)p_2$ . Since  $p_1 > 1/2$ ,

$$\mathbb{P}(B_1 \cap A(\mathcal{V}', \mathcal{V}'')) = p_1 p_2 > (1 - p_1)p_2 \geq \mathbb{P}(B_2 \cap A(\mathcal{V}', \mathcal{V}')),$$

so (7.7) is proved.

*Step 5.* Consider  $y_1, y_2 \in \mathcal{V}$  such that either (A1) or (A2) is satisfied and none of these vertices is the root. Let

$$(7.8) \quad \begin{aligned} B_1 &= \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_1, K_2 = y_2\}, & C_1 &= \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y_2\}, \\ B_2 &= \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_2, K_2 = y_1\}, & C_2 &= \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y_1\}, \\ B &= B_1 \cup B_2, & C &= C_1 \cup C_2. \end{aligned}$$

Consider the event  $A(\mathcal{V}', \mathcal{V}'')$  for some  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}(y_1, y_2)$  and suppose that  $v_* \in \mathcal{V}'$ . If we condition on  $A(\mathcal{V}', \mathcal{V}'')$ , the distribution of the order statistics of labels in  $\mathcal{V}'$  is independent of the values of the labels in  $\mathcal{V}''$ . Hence, (7.3) shows that conditional on  $A(\mathcal{V}', \mathcal{V}'')$ , the probability that the only peak in  $\mathcal{V}'$  is at  $v_*$  is at least  $d - 1$  times larger than the probability that it is at  $y_1$ . Since  $y_2$  is not adjacent to a vertex in  $\mathcal{V}'$ , if the only peak in  $\mathcal{V}''$  is at  $y_2$ , the only peak in  $\mathcal{V}'$  is at  $v_*$  and the highest label is in  $\mathcal{V}'$  then  $C_1$  holds. This and (7.7) imply that,

$$(7.9) \quad \mathbb{P}(B \mid A(\mathcal{V}', \mathcal{V}'')) \leq 2 \mathbb{P}(B_1 \mid A(\mathcal{V}', \mathcal{V}'')) \leq \frac{2}{d-1} \mathbb{P}(C_1 \mid A(\mathcal{V}', \mathcal{V}'')).$$

By symmetry, if  $v_* \in \mathcal{V}''$  then

$$(7.10) \quad \mathbb{P}(B \mid A(\mathcal{V}', \mathcal{V}'')) \leq 2 \mathbb{P}(B_2 \mid A(\mathcal{V}', \mathcal{V}'')) \leq \frac{2}{d-1} \mathbb{P}(C_2 \mid A(\mathcal{V}', \mathcal{V}'')).$$

It follows from (7.9)-(7.10) that for all  $(\mathcal{V}', \mathcal{V}'') \in \mathbb{V}(y_1, y_2)$ ,

$$(7.11) \quad \mathbb{P}(B \mid A(\mathcal{V}', \mathcal{V}'')) \leq \frac{2}{d-1} \mathbb{P}(C \mid A(\mathcal{V}', \mathcal{V}'')).$$

We will apply Step 1 with the family  $\{A(\mathcal{V}', \mathcal{V}'')\}_{(\mathcal{V}', \mathcal{V}'') \in \mathbb{V}(y_1, y_2)}$  (for fixed  $y_1$  and  $y_2$ ) playing the role of the family  $\{A_k\}$ . In view of (7.6), we see that (7.1) and (7.11) imply that

$$\mathbb{P}(B) \leq \frac{4}{d-1} \mathbb{P}(C).$$

Recall events  $C_1$  and  $C_2$  from the definition (7.8) of  $C$ . We have assumed that  $d \geq 10$ . The last estimate implies that for at least one  $i$  we must have

$$(7.12) \quad \mathbb{P}(C_i) \geq \frac{d-1}{8} \mathbb{P}(B) > \mathbb{P}(B) \geq \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_1, K_2 = y_2\}).$$

In preparation for proofs of other parts of the theorem, we present a stronger version of the last estimate under stronger assumptions. Suppose that for some  $m \geq 2$  we have  $\min_{i=1,2} \text{dist}(y_i, v_*) \geq m$ . Assume that  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}(y_1, y_2)$  and  $v_* \in \mathcal{V}'$ . If we condition on  $A(\mathcal{V}', \mathcal{V}'')$ , the distribution of the order statistics of labels in  $\mathcal{V}'$  (i.e., the joint distribution of the location of the highest label in  $\mathcal{V}'$ , the location of the second highest label in  $\mathcal{V}'$ , etc.) is independent of the values of the labels in  $\mathcal{V}'$ . Hence, (7.5) shows that conditional on  $A(\mathcal{V}', \mathcal{V}'')$ , the probability that the only peak in  $\mathcal{V}'$  is at  $v_*$  is at least  $(d-1)^{-m(m+1)/2}$  times larger than the probability that it is at  $y_1$ . If the only peak in  $\mathcal{V}''$  is at  $y_2$ , the only peak in  $\mathcal{V}'$  is at  $v_*$  and the highest label is in  $\mathcal{V}'$  then  $K_1 = v_*$ . We obtain using (7.7),

$$(7.13) \quad \mathbb{P}(B \mid A(\mathcal{V}', \mathcal{V}'')) \leq 2 \mathbb{P}(B_1 \mid A(\mathcal{V}', \mathcal{V}'')) \leq 2(d-1)^{-m(m+1)/2} \mathbb{P}(C_1 \mid A(\mathcal{V}', \mathcal{V}'')).$$

By symmetry, if  $v_* \in \mathcal{V}''$  then

$$(7.14) \quad \mathbb{P}(B \mid A(\mathcal{V}', \mathcal{V}'')) \leq 2 \mathbb{P}(B_2 \mid A(\mathcal{V}', \mathcal{V}'')) \leq 2(d-1)^{-m(m+1)/2} \mathbb{P}(C_2 \mid A(\mathcal{V}', \mathcal{V}'')).$$

It follows from (7.13)-(7.14) that for all  $\{\mathcal{V}', \mathcal{V}''\} \in \mathbb{V}(y_1, y_2)$ ,

$$(7.15) \quad \mathbb{P}(B \mid A(\mathcal{V}', \mathcal{V}'')) \leq 2(d-1)^{-m(m+1)/2} \mathbb{P}(C \mid A(\mathcal{V}', \mathcal{V}'')).$$

In view of (7.1), (7.6) and (7.15),

$$(7.16) \quad \mathbb{P}(B) \leq 4(d-1)^{-m(m+1)/2} \mathbb{P}(C) \leq 4(d-1)^{-m(m+1)/2} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2)).$$

*Step 6.* We will show that for any  $y$  whose distance from the root is greater than or equal to 2,

$$(7.17) \quad \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\}) > \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y, K_2 = v_*\}).$$

Let  $\mathcal{V}_i$ ,  $1 \leq i \leq d+1$ , be the subtrees obtained by removing the root from  $\mathcal{G}$ . Without loss of generality, we assume that  $y \in \mathcal{V}_1$ .

Let  $N = |\mathcal{V}|$ . A random labeling of  $\mathcal{G}$  can be generated in the following way. First, we generate independent random labelings  $L_j$  of  $\mathcal{V}_j$ , for  $1 \leq j \leq d+1$ . Then we assign a random label from the set  $[N]$  to  $v_*$  and divide randomly the labels remaining in the set  $[N]$  into  $d+1$  (unordered) families  $\Lambda_j$  of equal sizes, independently from  $L_j$ 's. Next,

labels in  $\Lambda_j$  are placed at vertices in  $\mathcal{V}_j$  in such a way that the order structure is the same as for  $L_j$ , for every  $j$ . Let  $L$  be the name of the resulting labeling of  $\mathcal{G}$ .

Let  $F_1$  denote the event that

- (i) For  $j \neq 1$ , the labeling  $L_j$  has only one peak and it is located at a vertex adjacent to  $v_*$ , and
- (ii)  $L_1$  has either only one peak at  $y$  or it has exactly two peaks, one at  $y$  and one at the neighbor of  $v_*$ .

It follows from Lemma 6.4 that

$$(7.18) \quad \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\} \subset F_1 \quad \text{and} \quad \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y, K_2 = v_*\} \subset F_1.$$

If  $L(v_*) = N$  then (i)-(ii) are not only necessary but also sufficient conditions for the event that there are exactly two peaks at  $v_*$  and  $y$ . More formally,

$$(7.19) \quad \{L(v_*) = N\} \cap F_1 = \{L(v_*) = N\} \cap \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\}.$$

If  $F_1$  holds,  $L(v_*) = N - j$  and  $L$  has exactly two peaks, with one of them at  $v_*$  and the other in  $\mathcal{V}_1$ , then all labels  $N, N-1, \dots, N-j+1$  must be in  $\mathcal{V}_1$ . Otherwise, because of (i), at least one of the numbers  $N, N-1, \dots, N-j+1$  would be the label of a vertex adjacent to  $v_*$  and, therefore,  $v_*$  would not be a peak. Given  $\{L(v_*) = N - j\} \cap F_1$ , the probability that all labels  $N, N-1, \dots, N-j+1$  are in  $\mathcal{V}_1$  is less than  $1/(d+1)^j$  for  $j > 1$  and equal to  $1/(d+1)$  for  $j = 1$ . We use (7.18) and (7.19) in the following calculation,

$$\begin{aligned} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y, K_2 = v_*\}) &= \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y, K_2 = v_*\} \cap F_1) \\ &= \sum_{j=1}^{N-1} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y, K_2 = v_*, L(v_*) = N - j\} \cap F_1) \\ &= \sum_{j=1}^{N-1} \mathbb{P} \left( \mathcal{N}(\mathcal{G}, 2) \cap F_1 \right. \\ &\quad \left. \cap \{K_1 = y, K_2 = v_*, L(v_*) = N - j, \{N, N-1, \dots, N-j+1\} \subset L(\mathcal{V}_1)\} \right) \\ &\leq \sum_{j=1}^{N-1} \mathbb{P}(\{L(v_*) = N - j, \{N, N-1, \dots, N-j+1\} \subset L(\mathcal{V}_1)\} \cap F_1) \\ &\leq \sum_{j=1}^{N-1} \frac{1}{(d+1)^j} \mathbb{P}(\{L(v_*) = N - j\} \cap F_1) \\ &= \sum_{j=1}^{N-1} \frac{1}{(d+1)^j} \mathbb{P}(\{L(v_*) = N\} \cap F_1) < \frac{1}{d} \mathbb{P}(\{L(v_*) = N\} \cap F_1) \\ &= \frac{1}{d} \mathbb{P}(\{L(v_*) = N\} \cap \mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\}) \\ &= \frac{1}{d} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\}). \end{aligned}$$

Since  $1/d < 1$  for  $d \geq 2$ , we conclude that (7.17) holds.

*Step 7.* It follows from (7.12) and (7.17) that  $\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_1, K_2 = y_2\})$  is maximized either when  $y_1 = v_*$  or if  $y_1$  and  $y_2$  are neighbors of  $v_*$ .

Assume that  $y_1$  and  $y_2$  are neighbors of  $v_*$  and  $y_3$  is a neighbor of  $y_2$ , but  $y_3 \neq v_*$ . We will show that

$$(7.20) \quad \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y_3\}) > \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_1, K_2 = y_2\}).$$

The proof will use the same ideas as the proof of Proposition 6.3. Recall the notation from Definition 6.1.

The probability  $\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y_3\})$  is the product of the following five factors, (i)-(v).

(i) The probability that the label of  $v_*$  is the highest among all labels. This probability is equal to  $1/|\mathcal{V}|$ , where  $\mathcal{V}$  is the vertex set of  $\mathcal{G}$ .

(ii) If we remove  $v_*$  from  $\mathcal{G}$ , we obtain  $d + 1$  subtrees. Let  $\mathcal{G}'$  be the subtree which contains  $y_3$ . Its vertex set will be denoted  $\mathcal{V}'$ . The second factor is the probability that  $y_3$  has the largest label in  $\mathcal{G}'$ . This probability is equal to  $1/|\mathcal{V}'|$ .

(iii) The third factor is the probability that the label of  $y_2$  is larger than the labels of all of its descendants in  $\mathcal{G}'$ , if we consider  $y_3$  to be the root of  $\mathcal{G}'$ . This probability is equal to  $1/\mathbf{n}_{y_3}^{\mathcal{G}'}(y_2)$ .

(iv) The fourth factor is the probability that the label of  $y_1$  is larger than the labels of all of its descendants in  $\mathcal{G}$ , with the usual root  $v_*$ . This probability is equal to  $1/\mathbf{n}_{v_*}^{\mathcal{G}}(y_1)$ .

(v) Just like in the proof of Proposition 6.3, we have to multiply by probabilities corresponding to all other vertices in  $\mathcal{G}$ , i.e., the last factor is  $\prod_{x \neq v_*, y_1, y_2, y_3} 1/\mathbf{n}_{v_*}^{\mathcal{G}}(x)$ .

The probability  $\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = y_1, K_2 = y_2\})$  is the product of the following five factors, (1)-(5).

(1) The probability that the label of  $y_1$  is the highest among all labels. This probability is equal to  $1/|\mathcal{V}|$ , where  $\mathcal{V}$  is the vertex set of  $\mathcal{G}$ .

(2) If we remove  $y_1$  from  $\mathcal{G}$ , we obtain  $d + 1$  subtrees. Let  $\mathcal{G}''$  be the subtree which contains  $y_2$ . Its vertex set will be denoted  $\mathcal{V}''$ . The second factor is the probability that  $y_2$  has the largest label in  $\mathcal{G}''$ . This probability is equal to  $1/|\mathcal{V}''|$ .

(3) The third factor is the probability that the label of  $v_*$  is larger than the labels of all of its descendants in  $\mathcal{G}''$ , if we consider  $y_2$  to be the root of  $\mathcal{G}''$ . This probability is equal to  $1/\mathbf{n}_{y_2}^{\mathcal{G}''}(v_*)$ .

(4) The fourth factor is the probability that the label of  $y_3$  is larger than the labels of all of its descendants in  $\mathcal{G}$ , with the usual root  $v_*$ . This probability is equal to  $1/\mathbf{n}_{v_*}^{\mathcal{G}}(y_3)$ .

(5) We have to multiply by probabilities corresponding to all other vertices in  $\mathcal{G}$ , i.e., the last factor is  $\prod_{x \neq v_*, y_1, y_2, y_3} 1/\mathbf{n}_{v_*}^{\mathcal{G}}(x)$ .

Note that the factors described in (i) and (1) are identical. The same applies to (v) and (5). Hence, it will suffice to show that

$$(7.21) \quad \frac{1}{|\mathcal{V}'|} \frac{1}{\mathbf{n}_{y_3}^{\mathcal{G}'}(y_2)} \frac{1}{\mathbf{n}_{v_*}^{\mathcal{G}}(y_1)} > \frac{1}{|\mathcal{V}''|} \frac{1}{\mathbf{n}_{y_2}^{\mathcal{G}''}(v_*)} \frac{1}{\mathbf{n}_{v_*}^{\mathcal{G}}(y_3)}.$$

The above inequality is equivalent to each of the following inequalities.

$$|\mathcal{V}'| \cdot \mathbf{n}_{y_3}^{\mathcal{G}'}(y_2) \cdot \mathbf{n}_{v_*}^{\mathcal{G}}(y_1) < |\mathcal{V}''| \cdot \mathbf{n}_{y_2}^{\mathcal{G}''}(v_*) \cdot \mathbf{n}_{v_*}^{\mathcal{G}}(y_3),$$



$$\begin{aligned}
& \frac{d^k - 1}{d - 1} \left( \frac{d^{k-1} - 1}{d - 1} \cdot (d - 1) + 1 \right) \frac{d^k - 1}{d - 1} \\
& < \left( \frac{d^k - 1}{d - 1} \cdot d + 1 \right) \left( \frac{d^k - 1}{d - 1} \cdot (d - 1) + 1 \right) \frac{d^{k-1} - 1}{d - 1}, \\
& \frac{(d^k - 1)d^{k-1}(d^k - 1)}{(d - 1)^2} < \frac{((d^k - 1)d + d - 1)d^k(d^{k-1} - 1)}{(d - 1)^2}, \\
& (d^k - 1)(d^k - 1) < (d^{k+1} - 1)d(d^{k-1} - 1),
\end{aligned}$$

$$(7.22) \quad -(d - 1)(d^k(d^k - d - 1) + 1) < 0.$$

It is easy to see that (7.22) is true for all  $d \geq 3$  and  $k \geq 2$ . It follows that (7.21) is also true. This completes the proof of (7.20) and, therefore, the proof of part (a) of the theorem.

(b) We have

$$\begin{aligned}
(7.23) \quad & \mathbb{P} \left( \mathcal{N}(\mathcal{G}, 2) \cap \left\{ \min_{i=1,2} \text{dist}(K_i, v_*) \geq m \right\} \right) \\
& = \frac{1}{2} \sum_{\substack{y_1, y_2 \\ \min_{i=1,2} \text{dist}(y_i, v_*) \geq m}} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = y_1, K_2 = y_2\} \cup \{K_1 = y_2, K_2 = y_1\})).
\end{aligned}$$

The factor  $1/2$  on the second line appears because we include both  $(y_1, y_2)$  and  $(y_2, y_1)$  in the sum.

The number of vertices in  $\mathcal{G}$  is  $(d + 1)(d^k - 1)/(d - 1) + 1$  so, assuming that  $d \geq 3$ , the number of pairs  $(y_1, y_2)$  is bounded by

$$(7.24) \quad \left( \frac{(d + 1)(d^k - 1)}{d - 1} + 1 \right) \left( \frac{(d + 1)(d^k - 1)}{d - 1} \right) \leq 4d^{2k}.$$

It follows from (7.16) that for  $m \geq 2$ ,

$$\begin{aligned}
& \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = y_1, K_2 = y_2\} \cup \{K_1 = y_2, K_2 = y_1\})) \\
& \leq 4(d - 1)^{-m(m+1)/2} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2)).
\end{aligned}$$

This, (7.23) and (7.24) imply that

$$(7.25) \quad \mathbb{P} \left( \mathcal{N}(\mathcal{G}, 2) \cap \left\{ \min_{i=1,2} \text{dist}(K_i, v_*) \geq m \right\} \right) \leq 2d^{2k} 4(d - 1)^{-m(m+1)/2} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2)).$$

If  $d \geq 3$  and  $k \leq \frac{1}{8}m^2$  then  $k \leq \frac{1}{4}m^2 \log(d - 1)/\log d$  and

$$\begin{aligned}
2d^{2k} 4(d - 1)^{-m(m+1)/2} & = 8 \exp(\log(d^{2k}(d - 1)^{-m(m+1)/2})) \\
& = 8 \exp(2k \log d - \frac{1}{2}m(m + 1) \log(d - 1)) \\
& \leq 8 \exp(-(m/2) \log(d - 1)).
\end{aligned}$$

We use this and (7.25) to see that if  $d \geq 3$ ,  $k \geq 2$  and  $m \geq \sqrt{8k}$  then

$$\mathbb{P} \left( \mathcal{N}(\mathcal{G}, 2) \cap \left\{ \min_{i=1,2} \text{dist}(K_i, v_*) \geq m \right\} \right) \leq 8 \exp(-(m/2) \log(d-1)) \mathbb{P}(\mathcal{N}(\mathcal{G}, 2)).$$

This proves part (b) of the theorem.

(c) The argument will be based on the comparison of  $\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\})$  and  $\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K_1 = v_*\})$ .

Suppose that  $y \in \mathcal{V}$  and  $m := \text{dist}(y, v_*) \geq 2$ . Let  $\Gamma$  be the geodesic between  $y$  and  $v_*$  and let  $z_j \in \Gamma$ ,  $0 \leq j \leq m$ , be such that  $\text{dist}(z_j, y) = j$ . Note that it is possible that  $z_2 = v_*$ . Suppose that  $\mathcal{V}'$  and  $\mathcal{V}''$  are obtained by removing the edge between  $z_1$  and  $z_2$  from  $\mathcal{G}$ ; we assume without loss of generality that  $v_* \in \mathcal{V}'$ . The corresponding graphs will be denoted  $\mathcal{G}'$  and  $\mathcal{G}''$ . If a labeling  $L$  of  $\mathcal{G}$  has the following properties: (i)  $L$  restricted to  $\mathcal{V}'$  has only one peak at  $v_*$ , (ii)  $L$  restricted to  $\mathcal{V}''$  has only one peak at  $y$ , and (iii) the largest label is in  $\mathcal{V}'$ , then  $L$  has exactly two peaks at  $v_*$  and  $y$ . Since  $v_* \in \mathcal{V}'$ , we have  $|\mathcal{V}'| > |\mathcal{V}''|$  and so the probability that the largest label is assigned to a vertex in  $\mathcal{V}'$  is greater than  $1/2$ . The distribution of the order statistics of  $L$  restricted to  $\mathcal{V}'$  is independent of this event, and the same holds for  $\mathcal{V}''$ . These observations and (6.1) imply that

$$\frac{\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\})}{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K_1 = v_*\})} \geq \frac{\frac{1}{2} \prod_{x \in \mathcal{V}'} (1/\mathbf{n}_{v_*}^{\mathcal{G}'}(x)) \prod_{x \in \mathcal{V}''} (1/\mathbf{n}_y^{\mathcal{G}''}(x))}{\prod_{x \in \mathcal{V}} (1/\mathbf{n}_{v_*}^{\mathcal{G}}(x))}.$$

Note that the factors corresponding to  $x \notin \Gamma$  are identical in the numerator and denominator, so

$$(7.26) \quad \frac{\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\})}{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K_1 = v_*\})} \geq \frac{1}{2} \prod_{j=2}^m \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_{v_*}^{\mathcal{G}'}(z_j)} \prod_{j=0}^1 \frac{\mathbf{n}_y^{\mathcal{G}}(z_j)}{\mathbf{n}_y^{\mathcal{G}''}(z_j)}.$$

Since  $m \leq k$ ,

$$(7.27) \quad \begin{aligned} \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_0)}{\mathbf{n}_y^{\mathcal{G}''}(z_0)} &= \mathbf{n}_{v_*}^{\mathcal{G}}(y) (\mathbf{n}_y^{\mathcal{G}''}(y))^{-1} = \frac{d(d^{k-m+1} - 1)}{d - 1} \left( \frac{d(d^{k-m+2} - 1)}{d - 1} \right)^{-1} \\ &= d^{-1} (1 - d^{-(k-m+1)}) (1 - d^{-(k-m+2)})^{-1} \geq d^{-1} (1 - d^{-(k-m+1)}) \geq \frac{d-1}{d^2}. \end{aligned}$$

The set of vertices corresponding to  $\mathbf{n}_{v_*}^{\mathcal{G}}(z_1)$  contains the set corresponding to  $\mathbf{n}_y^{\mathcal{G}''}(z_1)$  so we have the following bound for the second factor in (7.26) corresponding to  $\mathcal{V}''$ ,

$$(7.28) \quad \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_1)}{\mathbf{n}_y^{\mathcal{G}''}(z_1)} \geq 1.$$

Since  $\mathcal{V}' \subset \mathcal{V}$ ,  $\mathbf{n}_{v_*}^{\mathcal{G}}(z_j) \geq \mathbf{n}_{v_*}^{\mathcal{G}'}(z_j)$ , for all  $j$ . This implies that

$$(7.29) \quad \prod_{j=2}^m \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_{v_*}^{\mathcal{G}'}(z_j)} \geq 1.$$

We combine (7.26)-(7.29) to obtain for  $d \geq 3$ ,

$$(7.30) \quad \frac{\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\})}{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K_1 = v_*\})} \geq \frac{1}{2} \frac{d-1}{d^2} \geq \frac{1}{4d}.$$

We will now derive an upper bound. For  $1 \leq j \leq m-2$ , let  $\mathcal{V}_1^j$  and  $\mathcal{V}_2^j$  be obtained by removing the edge between  $z_j$  and  $z_{j+1}$  from  $\mathcal{G}$ ; we assume that  $v_* \in \mathcal{V}_1^j$ . The corresponding graphs will be called  $\mathcal{G}_1^j$  and  $\mathcal{G}_2^j$ . Recall the following notation,  $A(\mathcal{V}_1^j, \mathcal{V}_2^j) = \mathcal{N}(\mathcal{G}_1^j, 1) \cap \mathcal{N}(\mathcal{G}_2^j, 1)$ . Let  $K_1^{j,i}$  be the location of the highest peak in  $\mathcal{G}_i^j$ . It follows from the proof of Lemma 6.4 that

$$\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\} = \bigcup_{j=1}^{m-2} (A(\mathcal{V}_1^j, \mathcal{V}_2^j) \cap \{K_1^{j,1} = v_*, K_1^{j,2} = y\}).$$

These observations and (6.1) imply that

$$\frac{\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\})}{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K_1 = v_*\})} \leq \frac{\sum_{i=1}^{m-2} \left( \prod_{x \in \mathcal{V}_1^i} (1/\mathbf{n}_{v_*}^{\mathcal{G}_1^i}(x)) \prod_{x \in \mathcal{V}_2^i} (1/\mathbf{n}_y^{\mathcal{G}_2^i}(x)) \right)}{\prod_{x \in \mathcal{V}} (1/\mathbf{n}_{v_*}^{\mathcal{G}}(x))}.$$

The factors corresponding to  $x \notin \Gamma$  are identical in the numerator and denominator, so

$$(7.31) \quad \frac{\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\})}{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K_1 = v_*\})} \leq \sum_{i=1}^{m-2} \prod_{j=i+1}^m \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_{v_*}^{\mathcal{G}_1^i}(z_j)} \prod_{j=0}^i \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_y^{\mathcal{G}_2^i}(z_j)}.$$

For  $m \leq k$ ,  $d \geq 3$  and  $i \geq 1$ ,

$$\begin{aligned} \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_0)}{\mathbf{n}_y^{\mathcal{G}_2^i}(z_0)} &= \frac{d(d^{k-m+1} - 1)}{d-1} \left( \frac{d(d^{k-m+1+i} - 1)}{d-1} \right)^{-1} \\ &= d^{-i} (1 - d^{-(k-m+1)}) (1 - d^{-(k-m+1+i)})^{-1} \\ &\leq d^{-i} (1 - d^{-(k-m+1)}) (1 + 2d^{-(k-m+1+i)}) \leq 3d^{-i}. \end{aligned}$$

We estimate the other factors corresponding to  $\mathcal{V}_2^i$  as follows, for  $1 \leq j \leq i$ ,

$$\begin{aligned} \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_y^{\mathcal{G}_2^i}(z_j)} &= \frac{d(d^{k-m+1+j} - 1)}{d-1} \left( \frac{d(d^{k-m+1+i} - 1)}{d-1} - \frac{d(d^{k-m+j} - 1)}{d-1} \right)^{-1} \\ &= d^{j-i} (1 - d^{-(k-m+1+j)}) (1 - d^{j-i-1})^{-1} \leq d^{j-i} (1 + 2d^{j-i-1}) \leq 3d^{j-i}. \end{aligned}$$

Hence

$$(7.32) \quad \prod_{j=0}^i \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_y^{\mathcal{G}_2^i}(z_j)} \leq \prod_{j=0}^i 3d^{j-i} = 3^{i+1} d^{-i(i+1)/2}.$$

Next we deal with factors corresponding to  $\mathcal{V}_1^i$ , for  $i+1 \leq j \leq m$ ,

$$\begin{aligned} \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_{v_*}^{\mathcal{G}_1^i}(z_j)} &= \frac{d(d^{k-m+1+j} - 1)}{d-1} \left( \frac{d(d^{k-m+1+j} - 1)}{d-1} - \frac{d(d^{k-m+1+i} - 1)}{d-1} \right)^{-1} \\ &= (1 - d^{-(k-m+1+j)}) (1 - d^{i-j})^{-1} \leq 1 + 2d^{i-j}. \end{aligned}$$

This implies that

$$(7.33) \quad \prod_{j=i+1}^m \frac{\mathbf{n}_{v_*}^{\mathcal{G}}(z_j)}{\mathbf{n}_{v_*}^{\mathcal{G}_1}(z_j)} \leq \prod_{j=i+1}^m (1 + 2d^{i-j}) = \exp \left( \sum_{j=i+1}^m \log(1 + 2d^{i-j}) \right) \leq \exp \left( \sum_{j=i+1}^m 2d^{i-j} \right) \\ \leq \exp \left( \sum_{j=i+1}^{\infty} 2d^{i-j} \right) = \exp \left( \frac{2}{d-1} \right).$$

We combine (7.31)-(7.33) to obtain for  $d \geq 3$ ,

$$(7.34) \quad \frac{\mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap \{K_1 = v_*, K_2 = y\})}{\mathbb{P}(\mathcal{N}(\mathcal{G}, 1) \cap \{K_1 = v_*\})} \leq \sum_{i=1}^{m-2} \exp \left( \frac{2}{d-1} \right) 3^{i+1} d^{-i(i+1)/2} \\ \leq \sum_{i=1}^{\infty} \exp \left( \frac{2}{d-1} \right) 3^{i+1} d^{-i(i+1)/2} < \infty.$$

Since the lower and upper bounds in (7.30) and (7.34) do not depend on  $m = \text{dist}(y, v_*)$ , part (c) of the theorem follows.

(d) The number of  $y \in \mathcal{V}$  with  $\text{dist}(y, v_*) \geq 2$  is equal to  $(d^k - d)(d+1)/(d-1)$ . Since

$$d^k/2 \leq (d^k - d)(d+1)/(d-1) \leq 2d^k,$$

for  $d \geq 3$  and  $k \geq 2$ , part (c) of the lemma implies that for every  $d \geq 3$  there exists  $c_2$  such that for all  $k \geq 2$  and  $y \in \mathcal{V}$  with  $\text{dist}(y, v_*) \geq 2$ , we have

$$(7.35) \quad d^{-k}/c_2 \leq \mathbb{P}_2(K_1 = v_*, K_2 = y) \leq c_2 d^{-k}.$$

Consider  $y_1, y_2 \in \mathcal{V}$  and let  $m_i = \text{dist}(y_i, v_*)$  for  $i = 1, 2$ . Let  $m = \min(m_1, m_2)$ . We use (7.16) and (7.35) to see that

$$\begin{aligned} & \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = y_1, K_2 = y_2\} \cup \{K_1 = y_2, K_2 = y_1\})) \\ & \leq 4(d-1)^{-m(m+1)/2} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2) \cap (\{K_1 = v_*, K_2 = y_2\} \cup \{K_1 = v_*, K_2 = y_1\})) \\ & = 4(d-1)^{-m(m+1)/2} \mathbb{P}_2(\{K_1 = v_*, K_2 = y_2\} \cup \{K_1 = v_*, K_2 = y_1\}) \mathbb{P}(\mathcal{N}(\mathcal{G}, 2)) \\ & \leq 8(d-1)^{-m(m+1)/2} c_2 d^{-k} \mathbb{P}(\mathcal{N}(\mathcal{G}, 2)). \end{aligned}$$

This implies that for some  $c_3$  depending on  $d$  but not on  $k$ ,

$$(7.36) \quad \mathbb{P}_2(K_1 = y_1, \text{dist}(K_2, v_*) \leq m_1) = \sum_{j=0}^{m_1} \sum_{\substack{\{y_2: \text{dist}(y_2, v_*)=j, \\ \text{dist}(y_1, y_2) \geq 2\}}} \mathbb{P}_2(K_1 = y_1, K_2 = y_2) \\ \leq \sum_{j=0}^{m_1} (d+1)d^{j-1} 8(d-1)^{-j(j+1)/2} c_2 d^{-k} \leq c_3 d^{-k}.$$

A similar argument yields

$$(7.37) \quad \mathbb{P}_2(K_2 = y_2, \text{dist}(K_1, v_*) \leq m_2) \leq c_3 d^{-k}.$$

We use (7.36) and (7.37) to obtain

$$\begin{aligned}
(7.38) \quad & \mathbb{P}_2 \left( \max_{i=1,2} \text{dist}(K_i, v_*) \leq k - n/2 \right) \\
&= \sum_{j=0}^{k-n/2} \sum_{\{y_1: \text{dist}(y_1, v_*)=j\}} \mathbb{P}_2(K_1 = y_1, \text{dist}(K_2, v_*) \leq m_1) \\
&\quad + \sum_{j=0}^{k-n/2} \sum_{\{y_2: \text{dist}(y_2, v_*)=j\}} \mathbb{P}_2(K_2 = y_2, \text{dist}(K_1, v_*) \leq m_2) \\
&\leq \sum_{j=0}^{k-n/2} \sum_{\{y: \text{dist}(y, v_*)=j\}} 2c_3 d^{-k} \leq \left( 1 + \sum_{j=1}^{k-n/2} (d+1)d^{j-1} \right) 2c_3 d^{-k} \\
&= \left( 1 + (d+1) \frac{d^{k-n/2} - 1}{d-1} \right) 2c_3 d^{-k} \leq c_4 d^{-n/2}.
\end{aligned}$$

Suppose that  $|\text{dist}(K_1, K_2) - k| \geq n$ . It is possible that  $\min_{i=1,2} \text{dist}(K_i, v_*) \geq n/2$ . Suppose that the last condition is not satisfied and consider the case when  $\text{dist}(K_1, v_*) < n/2$ . If  $\text{dist}(K_2, v_*) > k - n/2$  then, by the triangle inequality,  $\text{dist}(K_1, K_2) \geq k - n$ , contradicting the assumption that  $|\text{dist}(K_1, K_2) - k| \geq n$ . It follows that  $\text{dist}(K_2, v_*) \leq k - n/2$ . Similarly, if  $\text{dist}(K_2, v_*) < n/2$  then  $\text{dist}(K_1, v_*) \leq k - n/2$ . This argument proves that

$$\begin{aligned}
& \{|\text{dist}(K_1, K_2) - k| \geq n\} \\
& \subset \left\{ \min_{i=1,2} \text{dist}(K_i, v_*) \geq n/2 \right\} \cup \left\{ \max_{i=1,2} \text{dist}(K_i, v_*) \leq k - n/2 \right\}.
\end{aligned}$$

It follows from the last formula, (7.38) and part (b) of the theorem that for every  $d \geq 3$  there exist  $c_4$  and  $c_5$  such that for  $k \geq 2$  and  $n \geq 2\sqrt{8k}$ ,

$$\begin{aligned}
\mathbb{P}_2(|\text{dist}(K_1, K_2) - k| \geq n) &\leq 8 \exp(-(n/4) \log(d-1)) + c_4 d^{-n/2} \\
&\leq 8 \exp(-(n/4) \log(d-1)) + c_4 \exp(-(n/2) \log d) \\
&\leq c_5 \exp(-(\sqrt{k}/64) \log d).
\end{aligned}$$

□

## 8. ACKNOWLEDGMENTS

We are grateful to Omer Angel, Jérémie Bettinelli, Sara Billey, Shuntao Chen, Ivan Corwin, Ted Cox, Nicolas Curien, Michael Damron, Dmitri Drusvyatskiy, Rick Durrett, Martin Hairer, Christopher Hoffman, Lerna Pehlivan, Doug Rizzolo, Bruce Sagan, Timo Seppalainen, Alexandre Stauffer, Wendelin Werner and Brent Werness for the most helpful advice. The first author is grateful to the Isaac Newton Institute for Mathematical Sciences, where this research was partly carried out, for the hospitality and support.

## REFERENCES

- [AB13] J. Ambjørn and T. G. Budd. Trees and spatial topology change in causal dynamical triangulations. *J. Phys. A*, 46(31):315201, 33, 2013.
- [Ale97] Kenneth S. Alexander. Approximation of subadditive functions and convergence rates in limiting-shape results. *Ann. Probab.*, 25(1):30–55, 1997.
- [BBPS15] Sara Billey, Krzysztof Burdzy, Soumik Pal, and Bruce E. Sagan. On meteors, earthworms and WIMPs. *Ann. Appl. Probab.*, 25(4):1729–1779, 2015.
- [BBS13] Sara Billey, Krzysztof Burdzy, and Bruce E. Sagan. Permutations with given peak set. *J. Integer Seq.*, 16(6):Article 13.6.1, 18, 2013.
- [BH90] Fred Buckley and Frank Harary. *Distance in graphs*. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, 1990.
- [CD81] J. Theodore Cox and Richard Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.*, 9(4):583–603, 1981.
- [DK14] M. Damron and N. Kubota. Gaussian concentration for the lower tail in first-passage percolation under low moments. *ArXiv 1406.3105*, 2014.
- [Ede61] Murray Eden. A two-dimensional growth process. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. IV*, pages 223–239. Univ. California Press, Berkeley, Calif., 1961.
- [HW65] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.
- [Qua12] Jeremy Quastel. Introduction to KPZ. In *Current developments in mathematics, 2011*, pages 125–194. Int. Press, Somerville, MA, 2012.
- [VdH14] R. Van der Hofstad. *Random Graphs and Complex Networks. Vol. I*. 2014. Book in preparation.

DEPARTMENT OF MATHEMATICS, BOX 354350, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195

*E-mail address:* burdzy@uw.edu

*E-mail address:* soumik@uw.edu